

DICAL ROOM  
AL LIBRARY  
OF MICH

III 8 - 1932

# REVIEW of EDUCATIONAL RESEARCH

---

Volume II

JUNE, 1932

Number 3

---

## TESTS OF PERSONALITY AND CHARACTER

---

AMERICAN EDUCATIONAL RESEARCH ASSOCIATION

*A Department of the*

NATIONAL EDUCATION ASSOCIATION

1201 SIXTEENTH STREET NORTHWEST

WASHINGTON, D. C.

## AMERICAN EDUCATIONAL RESEARCH ASSOCIATION

**T**HIS Association is composed of persons who are engaged in technical research in education, including directors of research in school systems, instructors in educational institutions, and research workers connected with private educational agencies. The Association became a Department of the National Education Association in July, 1930.

### Officers of the Association for 1932-33

#### President

WILLIAM S. GRAY  
University of Chicago

#### Vice-President

ALFRED D. SIMPSON  
New York State Department of Education

#### Secretary-Treasurer

WILLIAM G. CARR  
Research Division  
National Education Association

#### Executive Committee

The president, vice-president, and secretary, *ex officio*, and the following past presidents:

W. W. CHARTERS  
Ohio State University

J. L. STENQUIST  
Baltimore Public Schools

### EDITORIAL BOARD, 1932-33

FRANK N. FREEMAN, Chairman, University of Chicago  
J. CAYCE MORRISON, State Department, Albany, New York  
H. J. BAKER, Detroit Public Schools  
President and Secretary, *ex officio*

**Active Membership**—Persons eligible to membership in the Association must be recommended to the Executive Committee by a member who is in good standing. Upon approval of the recommendation, the person recommended will be invited by the Executive Committee to become a member of the Association. The Executive Committee has defined the qualifications for membership as follows:

"Membership in the Association is restricted to persons of good ability and sound training who are working in the field of educational research, and who can present satisfactory evidence in the form of published or unpublished studies which show ability to arrange, to organize, and to conduct research investigations and experiments. In addition, evidence of an abiding interest in the field of educational research is essential."

**Membership in the National Education Association** is a prerequisite to active membership in the American Educational Research Association. Any form of N. E. A. membership—annual, five-dollar, or life—satisfies the preliminary requirement.

---

Active members of the Association pay dues of \$5.00 per year. Of this amount, \$4.00 is for subscription to *THE REVIEW*. See back inside cover page of this issue. *THE REVIEW* is published in February, April, June, October, and December each year.

---

Entered as second-class matter April 10, 1931, at the post office at Washington, D. C., under the Act of August 24, 1912.

# REVIEW OF EDUCATIONAL RESEARCH

*Official Publication of the American Educational Research Association, a department of the National Education Association.*

---

Volume II

June, 1932

Number 3

---

## TESTS OF PERSONALITY AND CHARACTER

Part I of the report of the Committee on Character Tests and Psychological Tests: Katharine B. Greene, Paul R. Mort, Goodwin B. Watson, and Harry J. Baker, *Chairman*.

### TABLE OF CONTENTS

	<i>Page</i>
Foreword .....	184
Character Tests and Their Applications through 1930.....	185
GOODWIN B. WATSON, <i>Teachers College, Columbia University, New York, N. Y.</i>	
Bibliography .....	258

## FOREWORD

ON ACCOUNT of the large number of studies in the field of psychological tests the topic has been divided and will be treated in two numbers. The present number deals with tests of personality and character. The following number will deal with intelligence tests and tests of aptitude. Achievement tests, as originally planned, will be dealt with in a third number to appear during the winter.

The testing movement typifies, for most people, the scientific movement in education. This fact probably overemphasizes the importance of tests in comparison with the many other forms of scientific work in education. Studies of learning and of the psychology of the school subjects, for example, are fully as important as are tests. But the definiteness of tests and of the concept of individual differences which grew out of their use have impressed the minds of schoolmen and laymen alike. The result is that educators, within the space of fifteen or twenty years, have come almost universally to adopt the use of tests as a regular part of their procedure.

As was to be expected, some mistakes have been made in the use and interpretation of tests, due to over enthusiasm or to an incomplete knowledge of their nature and meaning. There are some indications that a reaction against the use of tests may set in. To offset the mistaken application of tests, and to prevent unintelligent reaction against them, it is important that full and authentic information concerning them be provided. To do this is the purpose of the three issues devoted to this subject.

FRANK N. FREEMAN, *Chairman, Editorial Board.*



## Character Tests and Their Applications Through 1930<sup>1</sup>

A REVIEW of character tests is unlike a review of tests of one aspect of personality such, for example, as intelligence. It is more nearly comparable with a review of all kinds of achievement tests, but covers an even wider range of concepts. To append a complete bibliography would call for approximately a thousand references and would fill the space allotted, in itself. The bibliography has, therefore, been confined in large part to bibliographies and to a few samples of each type of measure discussed. Certain phases of measurement which some might include with character and personality tests have here been omitted. Many performance tests, including those of Porteus and Ferguson, offer opportunity to study certain reactions of the individual to difficulty, but such tests have not been included here. Among the physiological indices of character the hundreds of studies upon the endocrine glands have been omitted from this review. Tests of pre-delinquency behavior are included, but the many studies on crime and delinquency, except as they involve some of the other types of tests being reviewed, have been left out. The enormous literature of case studies, while of immense value for the study of personality and character, has been omitted from this review as not specifically measurement.

The plan of the review will be the discussion, under each of the following headings, of the historical development of the test, the types of approach, the applications which have been made, and a list of published test blanks in cases where those are appropriate. Since the bibliography is very incomplete, many studies will be referred to only by name of author, the date of publication, and the bibliography in which the exact reference can be located. Thus, Starbuck (16, 1927) means an article by Starbuck, listed in the bibliography which is number 16 in the list at the end of this review, and appearing in 1927, i. e., May, Hartshorne and Welty's summary in the *Psychological Bulletin* for July, 1928. Often the same article appears in several bibliographies, but only one reference is given here, preference being given to the bibliography which offers an annotation, summary, or comment as well as a title. With the year and name of author it should also be possible to locate practically all of the recent references in the index number of *Psychological Abstracts*, and in the *Psychological Index*.

---

<sup>1</sup> Attention is called to the fact that the system of numbering references differs from that used in other issues of the *Review of Educational Research*. The reason for this deviation from general practice and an explanation of the form of reference used are given in the text.

## Headings under Which Character Tests Are Discussed

Abnormalities, Complexes, Symptoms of Maladjustment

Accuracy

Activity

Aesthetic Response

Aggressiveness, Dominance, Ascendancy, Submission

Appearance as an Index of Character

Confidence, Inferiority

Cooperation, Service, Negativism

Delinquent Trends

Emotions

Excitability

Expression, Handwriting, Will-Temperament

Happiness, Cheerfulness

Home Background

Honesty, Deception

Humor

Inhibition, Caution, Self-Control

Interests

Introversion, Extroversion

Leadership

Maturity: Social and Emotional

Moral Knowledge, Ethical Judgment

Morphology, Constitutional Type, Physical Build

Opinions, Attitudes, Prejudices

Originality, Imagination, Resourcefulness

Perseveration

Persistence, Perseverance, Effort

Physiological Indices of Character

Psychogalvanic Responses

Ratings, Reputation Measures

School Success and Failure: Character Factors

Self-Appraisal

Sex Differences

Sociability, Social Acceptability

Speed

Suggestibility

Types: Underlying Organization of Character

Summary

## Abnormalities, Complexes, Symptoms of Maladjustment

Almost thirty years ago Jung (40, 1905-06) proposed that the immediate association of individuals to standardized stimulus words, could be used to explore emotional complexes. Arnold (12, 1906) in a general discussion of association made somewhat similar proposals, and Yerkes and Berry (12, 1909) seem to have been among the first to publish results concerning the use of such a device with psychotics. Rusk (12, 1910) worked on the association reactions of children. The best known development along this line of measurement is the Kent-Rosanoff test (41, 1910; for test data see 21). Kent now feels that the instrument is practically worthless, but Rosanoff and many others give it a high rating. The test

early attracted the attention of Wells (12, 1911-12) who in 1911 and 1912 published several studies on the relation of reaction time, personal factors, practice effect, association types, response categories, and the like. The publication by Woodworth and Wells (52, 1911) is of classical importance in this field. The matter of reaction time was further studied by Crane (12, 1915) and with unusual thoroughness by Whately Smith (87, 1922). English (28, 1926) and some others have called attention to the discrepancy between reaction time as an index of a "complex" and peculiarity of the association as a complex indicator. The two indices are by no means invariably connected. On this and other grounds Sutherland (12, 1913) is critical of the association test technic. Kohs (12, 1914) and Hull and Lugoff (12, 1921) further developed the "complex indicators" showing that not only delayed or peculiar associations might be significant; but that certain other patterns, such as laughter, repeating the stimulus word, forgetting the stimulus word, answering too quickly, effect on the next words, and so on, might well be considered. Many investigators have used introspective analyses to help interpret the association responses. Typical of these studies is one by Burr and Geissler (12, 1913). A recent contribution to technic was made by Estabrook (30, 1930), who found that a specific suggestion (e. g., toward sex ideas) was not as potent in producing such responses as it is commonly supposed to be.

Norms for children were worked out by Woodrow and Lowell (51, 1916), although considerable experimenting with children's association had been done, of course, at earlier dates. Rosanoff (12, 1913), for example, reported on the application of the test to children. New adult norms for a rather different set of subjects were worked out by O'Connor (18, 1928). Still more recently experiments have been made with the association test as a group test by Elonen and Woodrow (18, 1928).

Association tests have been used as measures of affective disturbances by Henke and Eddy (12, 1909), Dooley (12, 1916), Tolman and Johnson (12, 1918), Griffiths (12, 1920) and many other investigators. They have also been used to explore sex differences by Haggerty (12, 1913), Bridges (16, 1927), and Miles and Terman (230, 1929); to study the relative strength of instincts by Moore (12, 1916), Allen (33, 1927), and Collman and McRae (33, 1927); egocentricity by Washburn (12, 1919) and Wells (12, 1919); cheerfulness by Washburn and others (102, 1919); associative inhibition by Kline (12, 1921); the measurement of will by Lewin (12, 1922); mood and performance by Sullivan (12, 1922); choice of salesmen by Freyd (28, 1926); originality by McClatchy (179, 1928); and differences between identical twins reared apart by Newman (18, 1929).

The next type of test to be evolved for discovering emotional maladjustments was the controlled answer questionnaire, in particular, the Woodworth Personal Data Sheet. (For data on test see 21.) This asked for "yes" or "no" answers to a series of very simple and direct questions involving symptoms which Woodworth found often mentioned in case

studies of psychoneurotics. This type of instrument demands, obviously, a high level of frankness and cooperation from the subject. One can make any sort of impression one wishes to make, by choosing the appropriate symptoms. On the whole, the impression made in the answers tends to be relatively consistent and constant. Fleming (34, 1928) reported a reliability of .89 with college freshmen, agreeing very well with the earlier results in the neighborhood of .90. Using the form for children, worked out by Mathews (43, 1923), Terman (138, 1925-30) found among his gifted children a reliability over a period of ten days of .75, and over five years a reliability of .42. The original Woodworth questions have been several times studied to discover the more and less diagnostic items. From the Emotional History Record (29, 1925), in which Chassell and Watson used many of the same questions, the twenty most diagnostic of general maladjustment were statistically chosen. Similarly Schneek (33, 1927) picked out the questions particularly diagnostic for epileptics and neurasthenics, but found that other types did not seem to have differentiating symptom-answers. House (34, 1928) found that government compensation men, drawing their income because of nervous upset consequent upon the war had, or thought they had, or pretended to have had very normal childhoods.

Laird (42, 1925; for data on test see 21) developed an important modification of the controlled-answer instrument. In his Colgate Mental Hygiene Scales, the answer is made not with "Yes" or "No," clearly a very crude and often ambiguous answer, but by a cross along a graphic rating scale between one extreme and the other. The "abnormal" answers on this scale were not determined *a priori*, but by marking off on the scale the space outside the middle 50 percent of student answers. Symptomatic responses were thus those which placed the individual in the extreme quartile of the group, without reference to the absolute meaning of the answer. Hoitsma (12, 1925) reported very satisfactory reliabilities for this technic.

The next extensive modification was made by Thurstone (48, 1930; for more data on test see 21), who went over all of the many questionnaires of this type and built an inventory somewhat more extensive than Woodworth's but containing many of the same questions. Thurstone's major contribution was in developing the evidence for internal consistency. He showed that each question differentiated in the direction of the test as a whole. Bernreuter (21, 1931) combined the Thurstone questions with others to make available in one instrument a measure of neurotic traits, of self-sufficiency, of introversion-extroversion, and of ascendance-submission. This is a practical modification of considerable clinical usefulness, but does not constitute an improvement in the theory of measurement.

Symonds and Jackson (30, 1930) improved methods for selecting mal-adjusted pupils by providing two questionnaires, one of which shows the pupil's impression of himself, the other, what his fellow pupils think of him. The second, or reputation test, is modelled on May's Guess Who Test,

which invites pupils to supply names of their comrades whom they think of as possessing the trait or characteristic in question. The questionnaire for the pupil himself is somewhat of an improvement over the Woodworth since it seems to ask more legitimate questions, flavored more by school work and less by the clinic.

Many other instruments, less well known, belong under this general heading of questionnaires designed to reveal abnormalities and complexes. Myerson (12, 1919) many years ago proposed a multiple-choice test which deserved, perhaps, more experimental attention than it has received. Heibredner (16, 1927) made use of a questionnaire for discovering the extent of inferiority feeling in pupils. This has been further developed for high-school pupils by Randolph Smith in a still unpublished dissertation. Faterson (30, 1930) used the inferiority indicator and also another scale, one registering worries. Cason (30, 1930) has for some years been collecting examples of annoying events, and from these has formulated a long list of annoyances. This makes a test suitable for distinguishing persons who report themselves much annoyed by many things, from those who report themselves little annoyed and by few things.

The relationship between tests like the Woodworth and intelligence has been a favorite item for study by such investigators as Marrow (34, 1928), Terman (138, 1925-30), Mathews (43, 1923), Witty (30, 1930), and Lamson (30, 1930); the results almost always point to a tendency for the brighter children to report fewer symptoms. This may mean that they have fewer problems, coming as they do from better homes and finding their school lot easier; or, it may mean, as Adler would say, that the lower intelligence score is a consequence of the emotional tangle; or, again, it may mean merely that the clever children are more discreet in their admissions. Probably each factor makes some contribution. Generally, however, the tests show a clear association of lower I. Q. with more symptoms, and greater suggestibility.

Questionnaires on symptoms, of one kind or another, have been used to study school failure by Goodrich and Clements (12, 1923), Young (33, 1927), Guthrie (16, 1927), Bridges (16, 1927), Peatman (34, 1928), Evans (30, 1930), Gilliland (30, 1930), McGeoch (30, 1930), and Fleming (34, 1928), with rather disappointing results. Rarely do school failures show more symptoms than other pupils of like intelligence but with better school records. Positive correlations run .03, .04, .11, and occasionally negative correlations are found.

These questionnaires applied to delinquents show, as a rule, emotional problems somewhat more numerous than in controls. Studies in this field are reported by Bridges (28, 1926; 16, 1927), Slawson (12, 1925), Cushing and Ruch (79, 1927), and Asher and Haven (30, 1930). Qualitative differences seem to be more pronounced than quantitative differences. Delinquents and criminals often show evidence of bad home conditions by report-



ing desire to run away from home, unhappiness in childhood, hatred of parent, fear of being left to go to sleep in the dark, and similar symptoms.

Other problems explored by means of symptom questionnaires include: race differences by Peatman (34, 1928), Garrett (18, 1929), and Sunne (12, 1925); stuttering by McDowell (34, 1928) and Dickinson (18, 1929); factors related to sex indulgence by Laird (28, 1926); to gambling by Hunter and Brunner (34, 1928); to success in camp leadership by Hendry and others (37, 1930); to success in Y. M. C. A. leadership by Sonquist (37, 1930); to choice of sports by Steen and Huntington (18, 1929); the study of characteristics of only children by Stuart (28, 1926); and the exploration of differences between identical twins raised in differing environments by Newman (18, 1929). Among the results of these studies none is more impressive than the evidence that stutterers, contrary to much psychiatric theory, appear to have no more symptoms of maladjustment, and, with the exception of the speech difficulty and its consequences, no different symptoms of maladjustment than may be found in equivalent groups of non-stutterers. The evidence that an only child is somewhat better adjusted than the child from a larger family may be misleading owing to the correlation between intelligence and the single-child family, while intelligence correlates negatively with the symptom questionnaires. For the most part, the attempt to use symptom questionnaires to select personnel, whether in education, business, or social agencies, breaks on the problem of frankness. Persons anxious to be selected do not paint themselves in unattractive colors. The few who are frank enough to do so, may show in that very answering, a distinguishing characteristic which compensates for the other handicaps.

Recognizing that frankness is not always easy to secure, numerous attempts have been made to construct instruments which would reveal emotional maladjustment without the subject's being aware that he was presenting such a picture. A decade ago Rorschach (46, 1921) published some painstaking analyses of the responses of normals and mentally disordered patients to a series of ten plates of paint-blot. He investigated the extent to which subjects answered in general or responded to details, the extent to which they were influenced by color as contrasted with form, the kinaesthetic imagery evident in their responses, the tendency to see people or animals in the misshapen figures, etc. His results were by no means conclusive, but the test has recently received some study by Beck (3, 1930) and much psychiatric ovation. Roemer (30, 1930) greatly improved the quantity and quality of analytic data, by using the Rorschach test in connection with accurate timing, stenographic reports, behavior observation during the testing, and a request that the patients draw the forms they seem to see.

The best known of the semi-disguised measures is the Pressey X-O Test (44, 1921; for data regarding test see also 21). This also appeared about ten years ago, and because of its simplicity of administration and inter-



esting content has a bibliography now of scores of titles. The test asks for the crossing out of items disliked, or worried about, or condemned. One section also explores associations in a multiple choice form. Two scores are offered, one for "affectivity," which shows the number of items to which the individual reacted, the other for "idiosyncrasy," commonly meaning the difference between the items to which the subject reacted and those to which the standardization group (unfortunately a very limited number of college students) responded. The authors have rather consistently maintained that the qualitative analysis of the direction of the aversions, etc. was much more useful than the numerical scores, but relatively few studies appear to have used the test in this manner. Chambers (12, 1925) used the same stimulus forms but scored them to indicate, in one case, emotional maturity, and in another case, probability of college success. These modifications have not been tried out on any groups other than the ones upon which they were developed; hence their general usefulness is not known. The reliability of the original test scores is somewhat variously reported, but the findings of McGeoch and Whitley (16, 1927), based on college sophomores, are fairly representative in showing the affectivity reliability between .5 and .8, the idiosyncrasy reliability somewhat less, .3 to .7. The test has been used in much the same fashion as the symptom questionnaires to study delinquents by Bridges (28, 1926; 16, 1927), Tjaden (16, 1926), and Branham (15, 1926); race differences, albeit with few cases, by Sunne (12, 1925) and Bond (28, 1926); characteristics of psychotics by Flowers; characteristics of criminals by Guilford (28, 1926) and Weber and Guilford (15, 1926); school success by Thompson and Remmers (34, 1928) and Fleming (34, 1926); characteristics of identical twins reared apart by Müller (28, 1925) and Newman (18, 1929); strength of instincts by Allen (33, 1927); and association with other tests of emotional qualities by Flügel (38, 1928), Gorham and Brotemarkle (18, 1929), Weber and Maijgren (18, 1929), and others named above. British norms were recently established by Collins (97, 1927). No psychological or educational generalization of merit has emerged from any of these studies. The results are usually negative, occasionally interesting in details, but so far have not been significant.

Some years ago Franzen (12, 1924) suggested that the relationship between the way in which an individual rated himself, the way he rated the average person, and an ideal person, offered interesting possibilities for studying personality. Watson followed this lead in constructing the Emotional History Record (29, 1925) and some of the Character Growth Tests for the Y. M. C. A. (15, 1926). This type of test at the college level was studied by Tyler (50, 1930), who found it unrelated to academic achievement, and by Sweet (47, 1929) in a very thorough study with boys twelve to fourteen years of age. The test measures some phases of self-criticism, of insight into other persons, of sense of superiority or inferiority, of appreciation of others, and of peculiarity in attitude and ideal. The

internal consistency of these measures is exceptionally high (.8 or .9) and Sweet (30, 1930) found some stability over several weeks. Their significance in relation to other indices of character has not been well established; although apparently boys with high insight have better reputations, boys with little deviation from the group do well on tests of moral knowledge, honesty, and cooperation.

It has been suggested by Goodenough and English that children's wishes might be of great diagnostic value. A test by Washburne, still in process of publication (21), uses this as one of its sections. Wishes are included in the diagnostic schedule created by Rogers (45, 1931). Rogers began with the material usually covered in psychiatric case studies of children, and during his experience in a child guidance clinic, formulated the procedure in a systematic outline. This served first as an individual oral interview and was then put into the form of a group test (21). The test sections are rather brief and ought to be more reliable, but they yield quantitative scores as well as valuable insights in relation to personal inferiority, social inferiority, family relationships, and day dreaming.

Travis (12, 1924) suggested a multiple-choice test which might be diagnostic of personality type, the alternatives presenting the sort of answer which might characterize one or another clinical entity. Town (34, 1928) described some eighty situations and observed the emotional and verbal response of the subject to these imaginative happenings. Schwartz presented a picture situation and classified the resulting responses as autistic, pleasurable somatic, adjustic, or fearful reactions. Loewey (34, 1928) suggested that some forms of infantile behavior can be rated in the response of the individual to his physical examination. Ball (18, 1929) offered an index of emotional instability in terms of behavior. Travis (49, 1926) described a laboratory test which has brought remarkably clear distinctions between schizoid and psychoneurotic types—the change of sensory threshold during reverie. Olson's book (58, 1929) is the most practical outline of a technic for the objective measurement of instability through nervous movements. It makes no claim that the nervousness measured by mouth movements, head scratching, and so on is closely related to problems of inner adjustment. Its report of an experiment with rats seems to point to a much more casual and superficial origin for such restless behavior. The book does, however, offer a first rate technic, based on many repeated short-interval observations, for recording the amount of external nervous activity.

The tests of abnormalities, complexes, symptoms of maladjustment have been discussed, first in terms of word-association technics, second in terms of symptom questionnaires, and third with reference to semi-disguised, disguised, and objective measures. It remains to be pointed out that some studies have not confined themselves to one type of measure but have made use of several, and have studied the intercorrelations, e. g., those of Allen (33, 1927), Flügel (38, 1928), Bridges (16, 1927), Weber and Maijgren

(18, 1929), Vernon (3), and Landis (86, 1925). As a rule, these are not high enough to warrant the interpretation that any one of these measures is a satisfactory index of the total area. To speak of them as tests of "emotional maladjustment" is justified only with the proviso that certain forms of maladjustment may be indicated by the test, but that certain others almost certainly are not. Moreover each measure seems, as yet, to be made up of many portions of error with relatively small portions of "pure" emotional-difficulty-intensity. The practical consequence is that until some more inclusive and less bulky battery is created, the study of individual or group adjustment can best proceed with a combination of many of these measures. Exception may be made for the Bernreuter, Colgate, Thurstone, Woodworth and Woodworth-Mathews tests which are derivatives from a common source and more or less interchangeable.

#### Test Materials Now Available<sup>1</sup>

<i>Title</i>	<i>Publisher</i>
Bernreuter Personality Record	Stanford University Press
Cason Annoyance Test	Stoelting
Chassell, Experience Variables	Chassell
Colgate Mental Hygiene Scales	Hamilton Republican
Kent-Rosanoff Association Test	Stoelting
Pressey X-O Test	Stoelting
Rogers, Emotional Diagnosis Test	Association Press
Rorschach Psychodiagnostik	Birchner
Schwartz, Social Situation Pictures	Stoelting
Sweet, Personal Attitudes of Younger Boys	Association Press
Thurstone, Neurotic Inventory	University of Chicago Press
Woodworth Personal Data Sheet	Stoelting
Woodworth-Mathews Test (for children)	Stoelting

<sup>1</sup> The alphabetical lists of tests available, following each section, refer to publishers by key name only. Full title and addresses are given below.

<b>Key Name</b>	<b>Addresses</b>
Association Press	347 Madison Ave., New York City
Birchner	Ernst Birchner, Berne, Switzerland
Bogardus	E. S. Bogardus, University of Southern California, Los Angeles, Calif.
C. E. I.	Character Education Inquiry, 129 E. 52 St., New York City
Center for Psychological Service	George Washington University, Washington, D. C.
Chassell	J. O. Chassell, University of Rochester Medical School, Rochester, N. Y.
Columbia University	Bureau of Publications, Teachers College, Columbia University, New York City
Hamilton Republican	Hamilton, N. Y.
Heidbreder	E. F. Heidbreder, University of Minnesota, Minneapolis, Minn.
Houghton Mifflin Co.	Boston, Mass.
Los Angeles Board of Education	Los Angeles, Calif.
MacNitt	R. D. MacNitt, State Teachers College, Superior, Wis.
Public School Publishing Co.	Bloomington, Ill.
Shields	F. J. Shields, Connecticut College for Women, New London, Conn.
Stanford University Press	Stanford University, Calif.
Stoelting	C. H. Stoelting Co., 424 N. Homan Ave., Chicago, Ill.
University of Chicago Press	University of Chicago, Chicago, Ill.
University of Iowa Press	University of Iowa, Iowa City, Ia.
World Book Co.	Yonkers-on-Hudson, N. Y.

## Accuracy

Accuracy has long been measured as a phase of almost every school and laboratory task. Arithmetic, spelling, tracing, letter cancellation, may all be considered types of the accuracy test. Washburne (33, 1927) proposed a measure of "consistency" in terms of the similarity between accuracy in one part of a test and accuracy in other parts. Presumably a test of relatively simple and uniform degree of difficulty should be used. Pollock (18, 1929) studied both accuracy and speed in a laboratory task which involved following a moving and variable pathway over a long time. Reputation for accuracy has found a place on rating scales for many years; typical examples are the study of clerical workers by the Bureau of Personnel Research of the Carnegie Institute of Technology (12, 1919), Dealey's study of problem children (12, 1923), Young's study of nurses (12, 1924), and Brandenburg's study of the personality and success of engineers (12, 1925). Hartmann's study (53, 1928) is perhaps the best and most inclusive. Accuracies in ability to tabulate, to follow directions, to estimate lengths, and to discriminate brightness, pitch, intensity, and weight, were measured. Reliabilities ranged from .96 for accuracy in estimating lengths to .39 for accuracy in following directions. Most of the intercorrelations were zero, indicating that no general unity in this trait can be assumed. Correlation of the total battery with another similar battery would be less than .30. The inevitable conclusion is that tests of accuracy must be defined in terms of the particular form and situation in which accuracy has been recorded.

Tests of accuracy, rather limited in scope in the light of the preceding observation, were used by Murray (12, 1920) in the vocational guidance of college women, by Bills (12, 1923) in determining efficiency at clerical work, by Hamilton (18, 1929) in studying the effect of incentives, and by Klineberg (54, 1927) in his very important observation that Indians respond to intelligence tests as they have learned to respond to their environment, with exceptional care and accuracy, but with no conception that speed is valuable.

## Activity

Measurement of the personalities of babies is confined necessarily to observation of some phases of their behavior. Excellent examples of the technics may be found in the study of the newborn by Pratt, Nelson, and Sun (59, 1930); in the study of the first year of life by Bühler, Hetzer, and Tudor-Hart (56, 1927); and in similar reports by Sherman (3, 1928) and Zoepfel (3, 1929). The most complete studies are being made by Gesell. Thomas (60, 1929) reported the attempt of several workers, notably Barker (55, 1930) and Loomis (57, 1931), to improve and standardize the technics of observation. Short periods are used; movements about the room, contacts with objects and persons are recorded in code. Correlation between one observer and another in the same situation may reach .98 or .99, a level not otherwise attained in character measures. Correlation of be-

havior in the same general setting on different days is high enough to allow this to be regarded as a consistent reaction of the personality. The extent to which behavior changes with change of situation has not yet been carefully studied. It would presuppose an analysis of the dynamic structure of the situation of the sort which Lewin (275, 1926-32) has been making. Goodenough (34, 1928) also worked with the short observation period in studying nursery school children, and found a consistency in the trait called *general activity* over several different days, as high as .8. Olson (58, 1929) used a similar technic for measuring such activity as lip movements, head scratching, hand-to-face, etc. during public-school study periods. He, too, found it possible by using twenty very brief periods of observation, and checking merely the presence or absence of the behavior, to secure reliabilities of .9 or better. The correlation which he found between nervous activity in one category, and nervous activity as measured by a different kind of movement, was .48.

General verbal activity, or talkativeness, as included in the Goodenough study, was observed among kindergarten children by Rugg, Krueger, and Sondergaard (18, 1929) and among other older children by Meltzer (14, 15, 1925). Wagner and Armstrong (3, 1928) studied a particular form of activity: the ability of children to dress themselves and to care for themselves.

### Aesthetic Response

In 1915 May (12, 1915) attempted to measure the reaction of individuals to various features of a service of worship. Watson (18, 1929) found that among adolescent boys stories, art, and music dominated over intellectual elements in producing the experience these boys regarded as worshipful.

Thorndike's early contribution (62, 1916) to the measurement of appreciation through visual forms appeared in 1916, and has been followed in recent years by a number of tests of artistic discrimination, notably those of Meier, the Los Angeles Test (61, 1926), the McAdory Test, and a still unpublished battery by Mendenhall. In the field of musical appreciation the Seashore Tests were pioneers, but these deal more with innate capacity, while the Kwalwasser-Ruch Test of Musical Information, Appreciation, and Accomplishment turns its attention more to the results of training.

Comparison of aesthetic with other values in the choice made by the individual was attempted by Watson (1926) in Forms E and F of the Summer Camp Tests used by the National Council of the Y. M. C. A., and by Allport in a test grounded on Spranger's "Lebensformen." The former covered so many areas as to be relatively unreliable; the latter shares the weakness of all self-report measures: that an individual can report any picture of himself which he chooses to give.

Tests of aesthetic responses have been used by Newcomb (12, 1924) to guide classroom practice, and by Smith (15, 1926) to study racial tastes.



A large number of studies of color preference have been omitted from this review.

#### Test Materials Now Available

<i>Title</i>	<i>Publisher</i>
Allport, A Study of Values	Houghton Mifflin Co.
Kwalwasser-Ruch Test of Musical Appreciation	World Book Co.
Los Angeles Test of Fundamental Abilities of Visual Art	Los Angeles Board of Education
McAdory Test of Artistic Discrimination	Columbia University
Meier Test of Artistic Aptitude	University of Iowa Press
Seashore Tests of Musical Aptitude*	Stoelting
Y. M. C. A. Summer Camp Tests, Form E and F	Association Press

#### Aggressiveness, Dominance, Ascendancy, Submission

In 1921 Moore and Gilliland (64, 1921) published a description of a series of tests for aggressiveness, including ability to carry on mental addition while subjected to distraction. Among the distractions used were electric shocks and a snake in near proximity, but the best results seemed to be obtained by the requirement that the subject look the examiner unfalteringly in the eye while working. Every waver counted against the subject, as did any loss in efficiency as compared with similar work in favorable isolation. Gilliland (28, 1926) later tried out the tests on 315 students, discarding the shock and snake, and adding the ratio of speeded to normal writing, a test borrowed from the Downey battery. The correlation of the test series with ratings was only .26. Freyd (28, 1926) found the tests useful in selecting successful salesmen.

The Allport A-S (Ascendancy-Submission) Scale (63, 1928) is a self-report device, asking for the subject's impression about his usual attitudes and practices. Thirty-five situations for women and thirty-three for men are included, each followed by a multiple-choice exercise. The reliability after a two-week interval is reported as .75, the correlation with ratings as .5 or .6, the higher figure corresponding to self-ratings. Correlations with intelligence, weight, height, family position, and scholarship are all close to zero. Correlation with a scale designed to measure extroversion was .38. Jersild (30, 1930), studying forty-two college students, found a correlation of about .5 between the Allport test and ratings.

Goodenough and Leahy (3, 1927) secured ratings on nursery-school children and found the oldest children in the family generally lacking in aggressiveness. Berne (195, 1930) used both tests and ratings, finding that social ascendancy correlated about .4 with mental age in pre-school children. Loomis' study (57, 1931) of social contacts included the differentiation between those contacts which were initiated by the individual himself, i.e., the aggressive ones, and those which were inflicted upon him by others, the passive contacts. Loomis' work demonstrated that these differ-



ences could be reliably determined; their significance in relation to the rest of the individual's living awaits further study.

Riddle (12, 1925) watched young men play poker, and analyzed the nature of the aggressive behavior shown in that situation.

#### Test Materials Now Available

<i>Title</i>	<i>Publisher</i>
Allport A-S Scale	Houghton Mifflin Co.
Bernreuter Personality Record	Stanford University Press

#### Appearance as an Index of Character

Among the early studies on the relation between appearance and character was one by Cogan, Conklin, and Hollingworth (12, 1915) later reviewed with many other studies in Hollingworth's book on *Judging Human Character* (67, 1922). About 1922 similar studies appeared by Paterson and Ludgate (12, 1922) on blondes and brunettes; by Dunlap (66, 1922) on the popular impressions about appearance indicating character; by Perrin (12, 1921) on attractiveness and repulsiveness; and by Pope (12, 1922) on the interpretation of the human face from photographs. The pioneer study of this latter type was Feleky's in 1914 (12, 1914). The conclusions of all of these investigations are well in accord: temporary emotional states can be identified from facial features, but the more permanent trends in character are not indicated by the measurement of any facial features. Later investigations by Buzby (12, 1924), Geissler (12, 1925), Cleeton and Knight (65, 1924), Winter (12, 1925), and Bender (34, 1928), all point in the same direction. Dunlap (33, 1927) showed that the mouth muscles rather than the muscles in the upper part of the face were the effective factors in influencing expression. Rice (28, 1926) showed that people do react to certain types of face in stereotyped form as "likely to be a bootlegger" or "likely to be a senator," albeit those reactions have no genuine correspondence in fact. Jersild (30, 1930) found that among college girls there was a correlation of .5 between ratings on beauty and ratings on amiability which might perhaps be explained by the easy recognition given to persons of attractive appearance. Wolff (30, 1929-30) found it difficult to match appearance and personality, even when the rigorous measurement of single aspects was abandoned, and the attempt made to get the impression or structure of the whole.

#### Confidence, Inferiority

Trow (70, 1923), Lund (69, 1926), Seward (17, 1928), and Jersild (18, 1929) studied confidence in terms of judgments and discrimination. Forms were used having certain previously learned features which markedly influenced the degree of confidence. Individual differences were found consistent with one sort of material, but Trow gave evidence to show that

confidence in one situation could not be considered an index of confident behavior in other situations.

A quite different approach was made by Heidbreder (68, 1927) in a self-report scale, including items supposed to reveal inferiority feeling. Faterson (30, 1930) found the reliability of this scale to be .73 after a six-week interval. Randolph Smith at the University of Minnesota High School developed a modification of this test in a form suitable for high-school students, which is not yet published. The Heidbreder scale was applied by Gardner and Pierce (18, 1929) to college students.

Degree of confidence in statements of opinion was measured by Williamson (12, 1915) and by Greene (18, 1929), the latter as an aspect of true-false tests. Cady (78, 1923) found differences in degree of confidence of raters in their judgment which are significant for the reliability of their ratings.

Bluffing, as studied by Fernberger (33, 1927) and by Thelin and Scott (34, 1928), may be viewed as a form of exaggerated self-confidence. Each found students quite willing to guess at answers to items to which they could not know the answer. The behavior studied seems, however, to be so greatly dependent on the unusual school-test situation that it is probably not of great significance in other realms.

### **Cooperation, Service, Negativism**

Although cooperation in some form almost invariably appears in rating scales, used for whatever purpose and at all ages, objective tests were late in appearing. Maller's work on the difference between the amount of work done by children when working for their own credit and the amount done when working for class credit is really one of the pioneer contributions (72, 1929). The Maller Test uses simple addition problems, although presumably any monotonous task would be appropriate. That it can hardly be regarded, without careful interpretation, as a measure of cooperation is shown by the work of some of Watson's students, still unpublished, indicating that for certain groups (e. g., boys against girls, our team against another team) pupils will work harder than for themselves; while for other groups, notably classroom units, they ordinarily do not respond with as much effort as they put forth for themselves. Thus this test must be regarded as a measure of cooperativeness of a certain kind in a certain situation. That this limited form of cooperation is well measured is indicated by Maller's self-correlation of .91. He found two-thirds of more than a thousand children working harder for self than for class. Almost all of the children were willing, however, on the "free choice" exercise at the end, to deed some of their work to count for the class rather than for themselves. Maller's finding, that the differences were negligible at the beginning of the period of work and increased as ennui entered, should be helpful in constructing other conduct tests along this line.

Hartshorne and May (71, 1929) added to Maller's tests measures of service in terms of willingness to help produce envelopes of jokes, pictures, etc. for the use of sick children, willingness to divide a cash gift with needy children abroad, and willingness to divide a school kit containing eraser, ruler, pencil-sharpener, etc. with children in another grade who had received none. Although these were the tests most widely used, there were other appeals given preliminary study and worth noting as stimuli to other investigators. Among these were coming early to school to work on material for hospital children, using shop time to make toy ducks for hospital children rather than continuing their own automobile project, and giving up an ice-cream dessert for charity. The low inter-correlations are again important. Correlation of ice-cream-giving with money-giving was .15. Correlation between both forms of giving on the one hand, and cottage mother ratings on usual helpfulness was -.03. Correlations among the five most used tests—money vote, school kits, hospital envelopes, Maller efficiency cooperation, and second Maller free choice—averaged .16.

Reputation for service was measured by: (1) a record of aid given to school projects; reliability, .80; (2) portrait matching, assigning one or another of ten sample sketches to each pupil; reliability of composite judgment from six teachers, using the scale twice, .84; (3) Guess Who Test, portraits matched by children with any of their classmates whom the portraits seemed to fit; reliability of ten service items by split halves, .88; (4) checklist of adjectives relating to helpfulness or its opposite, only those being checked for each child which the rater is sure are applicable; reliability of two forms, .48; and (5) conduct record, a series of multiple choice behavior descriptions, the teacher checking the one which most nearly fits each pupil. Total reputation for service had a reliability of .90; the inter-correlations among reputation measures averaged about .45. Reputation according to teacher agreed with reputation according to pupils to the extent of a correlation of .39. The correlation of the money-vote test with total reputation was about .17; the correlation of the school kits test with total reputation was .31; the correlation of the hospital envelopes with total reputation the same, .31; the Maller efficiency cooperation and the Maller free choice scores showed correlations of about .30 with total reputation. The entire battery of conduct tests agreed with the entire battery of reputation tests as indicated by a correlation of .52. Of the top quartile in test scores, 94 percent are rated above the neutral line for cooperation, while at the bottom quartile in test scores, 64 percent are rated below the neutral line.

Hartshorne and May (71, 1929) further studied the relation of service measures to other data, finding no significant relation (i. e., .40 or above) to age, intelligence, school grade, social status, acceleration, school marks, deportment marks, sex, physical fitness, emotional stability, general environment, occupational status, cultural level of home, economic level of home,

nationality of parents, religious affiliation, parental cooperation in filling out blanks, parental intelligence, length of attendance at school tested, or Sunday School attendance. Significant correlations appeared between siblings (.42) and between average boy and average girl of the class (.60).

Sorokin (34, 1928) studied the willingness of college students to contribute to their own department, and to other students in need at home and abroad. He found that only half who made altruistic professions agreed in their action. Zaluzhny (76, 1927) found increase in collective behavior with age, and with the influence of a cooperative environment. Children from peasant homes played individual games, children from the factory district played in larger groups. Berne (195, 1930) found an agreement as high as .76 between the ratings assigned young children and their obedient and cooperative behavior in test situations. Goodenough (18, 1929) and Lewy (15) reported on resistant or negativistic behavior as found in the test situation. This has been studied quantitatively by Rust (75, 1931) and Nelson (73, 1931), as reported in their respective dissertations. The most thorough study of non-cooperative or negativistic behavior was made by Reynolds (74, 1928) in the individual examination of 229 children aged 2 to 5. Correlation between one test and another was .2; correlation between one rating and another was .65. She found that negativism decreased with age, within the limits of the age group she used.

When boys in summer camps were asked by leaders to fill out some paper and pencil tests of no very great interest to the boys, Watson (111, 1928) used the number of questions unattempted as a measure of laziness or unwillingness to put forth effort for the camp program.

The influence of the cooperative situation was studied by Henning (34, 1928) in a valuable series of two-person experiments. He repeated various psycho-technical tests under conditions which showed how the individual reacts to working with one regarded as a rival, one regarded as a helper, one who works more rapidly or more slowly. The research included some eighty tests with twenty-five pieces of apparatus, and is of outstanding merit in opening up an almost unexplored field of undoubted significance for vocational and other adjustments.

#### Test Materials Now Available

<i>Title</i>	<i>Publisher</i>
Kits, Money, and Envelopes Tests	Association Press
Maller Efficiency Test, and Free Choice Test	Association Press

#### Delinquent Trends

The fact of arraignment before a court, and the further fact of being found guilty of some crime or misdemeanor, is of undoubted social significance. Its psychological significance is not so clear, but many studies

have endeavored to get at the factors which predispose to this unhappy outcome. Among the best are those of Glueck (80, 1930), Healy and Bronner (81, 1926), Slawson (12, 1924), and, best of all because of its careful matching with a control group of like age, sex, school grade, and neighborhood, that of Burt (77, 1925). The findings permit of considerable certainty that in a civilization like ours the delinquent will appear among the duller children, will be overage and uncomfortable at school, will have played truant, will come from a poor type of home, usually with one parent missing or incompetent, will fraternize with delinquent gangs.

In the attempt to get at the character elements in delinquent behavior more directly Cady (78, 1923) applied Voelker's tests on honesty, and a variety of other available measures of emotional stability and moral insight. The technic was advanced by Raubenheimer (12, 1925), who added tests of choice of companions, overstatement, recreational interests, etc. Terman and Laslett (12, 1925) showed how associations of delinquents led *bar* to mean saloon rather than candy, or *term* to mean jail rather than school. Along much that same line Schwesinger (28, 1926) found that knowledge of slang differentiated delinquents from other youngsters of similar intelligence. Association differences in delinquents had been earlier explored by Coriat (12, 1907) and Goddard (12, 1921). Cushing and Ruch (79, 1927) tried out on girls a battery of tests, similar to those used by Raubenheimer on boys, and found that delinquents could be differentiated from other girls of like intelligence, because they were more suggestible (12 P. E.), showed more symptoms of emotional instability (9 P. E.), made more over-statements on the false-book-titles test (5 P. E.), and showed poorer social attitudes in a paper and pencil opinion test (5 P. E.). The most thorough attempt to build a test diagnostic of probable delinquency was made by Lentz (82, 1925), who tried out most of the measures which up to 1925 had shown some promise. He found a number of considerable promise in differentiating between delinquents and well-behaved youngsters of similar intelligence and home background. Lentz, however, took his tests one more step. He tried them out on a second pair of groups, to see whether the items which proved useful the first time also held for the second comparison. Unfortunately they did not. Most of the promising items dropped out. The two which seemed most useful were a questionnaire on activities and interests, and a "daily contribution test," which called for bringing in each day some little item of interest. The numerous other studies in which one or two of the more popular tests are administered to a group of delinquents and the results compared with published norms, need not be reported.

### Emotions

Some dozen different approaches have been used in the measurement of emotions: facial or vocal expression, word association, psychogalvanic



response, heart rate, blood pressure, breathing curve, oxygen consumption or basal metabolism, body chemistry changes, interference with other work, self-report, and observation. Outstanding in the early analysis of facial expression is the work of Feleky (12, 1914), Langfeld (12, 1918-19), Ruckmick (12, 1921), Pope (12, 1922), and Dunlap (33, 1927). Sherman (16, 1927) showed that the judgment of emotional responses of infants when portrayed in motion pictures, or heard to cry from behind a screen which kept the stimulus situation invisible, was rather precarious. Unless the observer knew the stimulus situation it was not possible clearly to separate the emotional responses of fear, anger, pain, and the like. It would seem, therefore, that the early behavioristic accounts involved considerable interpretation. Landis (18, 1929) gave further data on facial expression in emotion, analyzed by means of special marking of the facial muscles before photographing. Association tests have been discussed under a separate heading. Here mention should be made of Moore's (12, 1916-17) word association methods for testing anger, fear, sex, and other "instinctive" trends in the individual. Smith's book (87, 1922) summarizes most of the association and galvanometer technics. The psychogalvanic studies are summarized in a later section of this review.

Pulse changes in emotion were analyzed by Landis and Slight (18, 1929); blood pressure changes in response to vivid moving picture scenes by Marston (12, 1924), Nisson (34, 1928), and Scott (30, 1930). Chemical changes during emotion point to an increase in protein and sugar indices of the urine during such an emotional strain as an examination (15, 1926). The respiratory changes in children were discussed by English (15, 1926), while Roemer (30, 1930) reported an apparatus so light that it can be worn during active exercise, and of such a nature that it does not interfere with free bodily movements.

Oxygen consumption and basal metabolic changes during emotion have been the subjects of studies by Totten (15, 1925), who found an increase of from 5 percent to 25 percent due to emotional response, by Henry (18, 1929), and by Segal, Binswanger, and Strouse (34, 1928), who concluded that the emotion connected with a threatened operation did not effect basal metabolism unless there was an accompanying thyroid disorder.

It has long been observed that steadiness and concentration on intellectual tasks are disturbed by strong emotions, but this fact has not often been used as a test. Watson (29, 1927) presented theological students with printed material in which nouns were to be crossed out. No noticeable differences in rate were observed to correspond to material of a humorous, abstract, sentimentally religious, or anti-religious character, but material from a very modern treatise on sex relations produced a marked decrease in the efficiency of work.

Diaries in which students recorded emotional experiences, together with certain facts about the cause and duration, were analyzed by Gates (28, 1926) and by Flügel (15, 1925). Direct observation of the emotional



behavior of children was the basis of the work of Jones (34, 1928), of Goodenough, who found that anger could be observed during a series of thirty second periods with a reliability of .6 (34, 1928), and of Herring (84, 1930) and Gauger (83, 1929), who found it possible to record the reaction of children to various tastes, with extremely high reliabilities.

Many of the best studies of emotion have involved comparison of a battery of measures of various types. Stratton (88, 1926) used diaries and reaction to imaginary situations. Skaggs (28, 1926; 30, 1930) found that startle-emotions accelerated breathing, retarded the heart, and disturbed steadiness, but that these factors were considerably influenced by the "set" of the individual. Much of the best work was done by Landis (86, 1925; see also 12, 1924; 28, 1926), who exhibited great ingenuity in creating emotions in subjects (picking wet frog from pail, beheading a white rat, etc.) and recorded facial, heart, respiratory, and galvanic responses simultaneously. It would appear that the various measures of emotion do not always agree. Patterns are more apt to be consistent within the individual, but not apt to be consistent from one person to another. Such questionnaires as the Woodworth-Mathews measure something rather different from emotionality as shown in specific situations.

### Excitability

Closely related to emotional responsiveness is the factor of irritability (276, 1928) or ready excitability. The emotional responses to a series of situations constitute, of course, one approach to the measurement of this aspect of personality. Washburn and others (28, 1926) used the recall of previously experienced emotions to differentiate between calm and emotional individuals. Cason's (90, 1930) long list of items which some people have found annoying has been made into a test, with four degrees of response to each item. Like other self-report measures it is presumably sensitive to the impression the subject wishes to create. Hewlett and Lester (34, 1928) rated subjects on expressiveness during a standardized interview. Correlations with intelligence and with self-rating on introversion were zero.

Physiological indices seem more clearly related to this than to any other character trait. Rich's (91, 1928) correlations are low, but suggest that the calm individual has more acidity in saliva and urine, with higher creatinine in these body fluids. Mateer (28, 1926) and later Timme in a report at the First International Neurological Congress, presented strong evidence for believing that calcium inadequacy is closely related to hyper-irritability in children.

#### Test Materials Now Available

Title	Publisher
Cason Annoyance Test	Stoelting

## Expression, Handwriting, Will-Temperament

For a generation or more there has been scientific as well as popular interest in the hypothesis that the individual is so characteristically a unit that every movement and expression has something about it which identifies it uniquely with him. Handwriting has been prominent in the movements studied, one of the early studies having been made in 1906 by Binet (92, 1906). Hull and Montgomery (94, 1919) in this country made some attempt to check up on the validity of the claims that certain specific elements in writing, regardless of accompaniments, were associated with certain character traits as measured by ratings. Such analyses have not substantiated the claims, but the claims have altered. Klages (95, 1920), in the best known book on handwriting and character, built a total philosophy of personality, within which general forms and relationships must be taken into account. These claims are best checked by matching experiments, in which the total character of the person and of the handwriting can be studied in its natural structure. Unfortunately the methods for building such character criteria are seriously inadequate. Kinder (28, 1926) and Newhall (23, 1926) found that sex of the writer could be determined by untrained judges 60 percent to 70 percent of the time. Krauss (30, 1930) found that attempts to draw lines symbolic of emotional states gave products which could be matched correctly, e. g., anger, or reverie, in 70 percent of the cases. Arnheim (3, 1928) and Wolff (99, 1930) matched personality sketches and handwriting, the latter finding success about twice as often as would have been expected from chance alone.

Handwriting has been used to study characteristics of the insane by Barillot (12, 1922), and of prospective employees by French (12, 1917) and Hollingworth (67, 1922), as well as in pure characterology. A related line of evidence is emphasized by Lembke (30, 1930), who found that the drawings of aggressive pupils could be differentiated from the shy ones.

Duffy (30, 1930) studied the muscular tension in the grip of children making multiple-choice reactions, and found the shape of the curve, as well as its level, related to characteristics of excitability and work habits as described by teachers.

The Downey Will-Temperament Tests (93, 1923), an outgrowth of Downey's interest in graphology as an expression of personality (12, 1919), represent a bibliography of 77 titles at hand for the preparation of this review. The complete list is assuredly longer. The extraordinary amount of attention given to a test which very quickly showed itself to be unreliable and unrelated to the popular understanding of the trait names if used, can be explained in part in terms of the unique character of the test, and in part by its pioneer appearance. It still remains the stock illustration of character testing in most psychological texts. A tentative scale appeared in 1912 (12, 1912), but most of the use of the test began about 1920. A

group form appeared in 1919, an individual form in 1921, Ream's form in 1922, and a non-verbal form in 1927; the non-verbal showed an average correlation with the verbal of .24. Criticisms, partly from the nature of the tasks set, but largely on the basis of very low reliabilities and complete lack of accord with other data about individuals were made by Ruch (12, 1921), Filter (12, 1921), Meier (12, 1923), Ruch and DelManzo (12, 1923), Herskovits (12, 1924), Hurlock (28, 1926), Stoddard and Ruch (28, 1926), Ruch and Manson (15, 1926), Downey and Uhrbrock (16, 1927), and Gorham and Brotemarkle (18, 1929). Attempts made by these writers included matching with ratings, matching with self-ratings, and effort of the individual to pick out his own profile, but none of these methods gave any support to the validity of the terminology. After May's (96, 1925) rather devastating summary of the evidence, Uhrbrock and Downey reported further applications of the test to college women, with reliabilities between .31 and .63; to junior high-school pupils with reliabilities between .09 and .64. The best review is in Uhrbrock's dissertation (97, 1928), which shows self-correlations for the group test from .12 to .89, with an average of .52; for the Carnegie adaptation from .20 to .83, with an average of .54; for the non-verbal from .24 to .71, with an average of .45 and a cross-correlation value of .20. Relations to ratings averaged .03, to school grades .03, to intelligence .08.

Attempts to predict scholarship were made by Stone (12, 1922), Poffenberger and Carpenter (12, 1924), Miner (12, 1925), Traxler (12, 1925), Flemming (15, 1926), Downey (16, 1927), Kornhauser (16, 1927), and Oates (34, 1928). The results of the last named study, showing that the closest correlations were found between scholarship and speed (.26, .11), scholarship and aggressiveness (-.08, .26), scholarship and confidence (.30, .00), are typical. Miner tried the neat experiment of sending to the test author the profiles of some pupils who were in scholarship below their intelligence expectation, others who were superior in actual achievement. The profiles were sorted with the 50 percent accuracy which could have been obtained by chance.

The tests were applied to delinquents by Bryant (12, 1921), Clark (12, 1921), Wires (15, 1926), Branham (15, 1926), and Downey (33, 1927), resulting in no clear distinctions from normals. They were applied to small groups to determine race differences in studies by MacFadden and Dashiell (12, 1923), Sunne (12, 1925), Bond (28, 1926), Garth and Barnard (16, 1927), and Peterson and Lanier (17, 1929), with no important and consistent distinctions. They were applied in the selection of teachers by Kolstead (12, 1924) and Thompson (34, 1928), of salesmen by Ream (12, 1921) and Freyd (12, 1922-24), of successful dentists by Roe and Brown (16, 1927), and of personality types by Downey (12, 1924) and Oates (18, 1929), but in no case with valuable results. They have been applied in relation to morphology by Naccarati and Garrett (12, 1924), to individual guidance by Reaves (12, 1925), to speech prob-

lems by Wagoner and Downey (98, 1922), to differences between identical twins by Müller (15, 1926) and Newman (18, 1929), to variability of traits within the individual by Hull (33, 1927), and to the study of adaptation to institutional life by Harrel and Davis (18, 1929), with steady disregard of the unreliability of individual scores. The Downey Test is one of the few American tests of personality and character used in England by Collins (12, 1925), Richardson (18, 1929), and Oates (3, 1929).

#### Test Materials Now Available

<i>Title</i>	<i>Publisher</i>
Downey Group Will-Temperament Tests	World Book Co.
Downey Individual Will-Temperament Tests	World Book Co.

### Happiness, Cheerfulness

It is certainly quite as important that an individual's way of living be satisfactory to himself as that it should be useful to his fellows, but this inner aspect of adjustment has not been so extensively studied. Washburn and others (102, 1919; 12, 1919; 28, 1926) explored some tests to differentiate the girls of generally cheerful mood from those of generally depressed mood. The cheerful girls when asked to recall emotional experiences, recall a larger proportion of pleasantly toned experiences; their word associations lead more directly to pleasant associations.

Hamilton (101, 1929) measured the happiness of two hundred married persons in their conjugal relationships by the answers to fourteen questions, given in carefully standardized oral interview and under conditions of voluntary participation. Watson (37, 1929) studied the life satisfaction of adolescent boys, using questions selected from a test of the Woodworth Questionnaire type. In a later study Watson (104, 1930) applied five different devices for self-report of happiness: mark along a graphic rating scale for general happiness, choice of one of twelve suggested descriptions of prevailing mood, original description of mood scored by judges, list of optimistic and pessimistic adjectives, with directions to check those which usually applied to self, and average of graphic scale self-rating for happiness in such areas as health, vocation, sex life, friendships, hobbies, and religion. Differentiating extremely well-satisfied, average, and extremely unhappy groups, an attempt was made to find the factors in previous experience which might have contributed to this condition. Sailer (103, 1931) modified Watson's measure and applied it to young men in industry. The reliability of the happiness indices by split-halves appears to be about .85, consistency over several weeks about .60. The measure is useful, obviously, only under conditions in which frankness can be guaranteed. Hersey (18, 1929) attempted to measure morale and general feeling tone in industry, and developed incidentally an excellent technic for assuring the participants

of anonymity in their answers, allowing them to pass out blanks, shuffle them around, collect them, mark with underlining only, and so on. Fairchild (30, 1930) studied men in the metal trades and found their happiness, as best he could estimate it, closely related to their degree of skill. Jasper (18, 1929) used a multiple-choice test which gave the subject a chance to choose a predominantly optimistic or pessimistic conclusion regarding each of the issues presented.

A third and the most objective method of approach to this very subjective state, is by behavior observation. Thomas and Gregg (60, 1929), Enders (100, 1928), Washburn (18, 1929), and others observed smiling and laughing in young children. This may be a better index of their relative joy in living at that age than it would be if applied to older individuals.

#### Test Materials Now Available

<i>Title</i>	<i>Publisher</i>
Happiness Report	Association Press

#### Home Background

Home background may not seem like a character trait, but in the light of a number of studies it is a better index of character than many tests which deal with the supposed result rather than with so basic a causal influence. Barr (12, 1921) proposed a social rating scale which is still widely used. The Whittier Scales for rating homes and neighborhoods are now superseded and out of print. The Sims Score Card (105, 1925) does very well for economic status in small American cities but does not fit unusual social environments. Burdick's Scale gets at the matter indirectly through the child's concept of what belongs in a living room, of how fathers speak to children, of what constitutes good manners. This scale touches cultural influences rather more than the others, but has a correlation of .4 to .7 with occupational status, with the Sims Scale, and with ratings by home visitors. The recently published McCormick Scale has a high reported reliability and is more comprehensive in its estimate of family life than any of the others. Wylie (33, 1927) found that about 90 percent of the answers given by children on a "facts about the home" questionnaire were truthfully given. Hartshorne and May made extensive use of home background measures in the Character Education Inquiry (108, 1928; 71, 1929; 146, 1930) and found correlations of about .30 between home background and honesty; about .20 between home background and each of the other conduct factors: service, persistence, and inhibition. Moral knowledge showed a fairly high correlation (54) with the Burdick Test but less relation (24) to the Sims Score Card.



## Test Materials Now Available

<i>Title</i>	<i>Publisher</i>
Burdick Apperception Test	Association Press
McCormick Scale for Measuring Social Adequacy	Stoelting
Sims Score Card for Socio-Economic Status	Public School Publishing Co.

### Honesty, Deception

It has been consistently recognized by the advocates of character testing that the approach to honesty, in which deception was, in the very nature of the case, to be expected, could be made only through situations in which the subject did not know the purpose of the enterprise. The pioneer study was made by Voelker (110, 1921), the originality of whose contribution remains unsurpassed in the realm of character testing; although his application of his tests was unfortunately inadequate, due to groups unequated for age or mental age, and to unreliable test batteries. Voelker's tests, the most popular of which have been the peeping test (trying to draw some design or check some forms, with closed eyes), the waxed paper to preserve the original school-work record so that changes made to raise the score could be counted, and the overstatement of knowledge or ability type of test, have been applied and slightly modified in further studies by Perry (12, 1923), Cady (78, 1923), Raubenheimer (12, 1925), Terman (138, 1925-30), and others.

The largest and most significant enterprise to date in honesty-measurement, is the work of Hartshorne and May (108, 1928) in the Character Education Inquiry. Their book reviews a score of previous attempts to measure deception and develops group tests, applicable in school rooms on a large scale. They used ten tests of copying from a key or answer sheet, six of adding on more scores after time has been called, three of peeping when the eyes should be closed, five of faking a solution to a too-difficult puzzle, four of faking a score in a physical ability contest, one of getting help on a test which the individual had promised to do alone, one of exaggerating one's own virtues, and several representing minor forms of cheating in parlor games. Reliabilities range from .24 to .87; most of them are above .7. The proportion of dishonest behavior varied with the type of test. Some 20 percent made gross exaggerations of their own virtues, some 80 percent peeked on at least one test. Correlations with other measured factors were below .40 except for the relation of honesty to suggestibility (-.60), to intelligence (.40 in some tests, not all), to good home-culture background (.40 in some tests), and agreement of behaviors of friends in same school class (82, 1925). Comparison of consistently honest and less consistently dishonest children showed the honest group to come from more favored race and nationality, to be of higher intelligence, and to



have a better home background. The usual educational agencies, such as school lessons on honesty, religious training, and character building clubs and camps, seemed to have no consistent effect upon the honest behavior of participants. Cheating was, however, definitely less marked in the "free" type of school, as contrasted with formal schools serving a like group of children. Correlations between honesty on these tests and results from other tests were as follows: with service .3, with persistence .1, with inhibition .3, with moral knowledge .4, with reputation .2. The average intercorrelation between one test of honesty and another was no higher (.2) than these correlations across trait lines, suggesting that there is not much more basis for calling one of these tests a measure of the total trait of honesty than there is for calling it a measure of any other desirable character quality.

This may be a good opportunity to mention some of the new technics introduced into character testing, thanks to the work of Hartshorne and May. Their method of validating tests by computing the correlation between the tests given and in infinite series of tests of which those given are only a random sample, is very helpful; but it must be borne in mind that any characteristic common to the sample (e. g., schoolroom administration) must be thought of as common to the infinite series criterion. Hence it is somewhat misleading to suggest that a fifth, or any other fraction of the total area involved in a trait like honesty or service, has been measured by the given battery, unless that battery represents a sampling of all situations in which the trait could be manifested, without irrelevant constant concomitants. Using this technic Hartshorne and May further predicted the number of tests needed for a correlation of .90 to .95 with a criterion made up of an infinite number of such tests. As a rule, some thirty conduct tests were seen to be needed, even when the sphere of generalization is limited to tests administered to classroom groups under school conditions. To measure any trait in all its ramifications would require many more such units, for the intercorrelations fall rapidly with change of situation. The folly of generalization about individual character traits from a single conduct test, or even from a battery of half a dozen, is apparent.

The division of the standard deviation of means by the average standard deviation of the mean for the groups studied (average quotient 3.0) forms an interesting technic for showing the extent to which results are influenced by factors peculiar to the group. Hartshorne and May usually called this factor group morale, but it is probable that there was in it also a very large element of "errors-common-to-the-administration-to-this-class-at-this-time," since the class units were also the units tested.

The choice of groups in three different types of community, and the contrasting results, correlations not infrequently changing in sign with difference in social structure, ought to serve as a needed correction on practically all previous (and most later) test result publication. The factors influencing these behavior results are so largely found in the structure of the community, that results published without some analysis and description

of the community in which they were obtained are as ill-controlled as would be results published on groups whose age, sex, race, and intelligence were unknown.

The combination of all known data about the individual character, after a long period of varied types of measurement, the sorting of these portraits along an imaginary scale of "general all around character," gave another interesting technic. On that basis the school honesty data had a correlation of about .5 with the total evaluated character. It stood about midway between items like reputation measures (.7) and mental age or emotional stability (.3).

The use of consistency in the behavior tested, as well as average level of that behavior, was another important contribution. *Integration*, the term used by Hartshorne and May, seems unfortunate because it suggests a functional organization which they did not study. As a rule they found that numerous children were consistently good (i. e., honest, helpful, etc.), but that very few were consistently and all the time bad (i. e., deceptive, selfish, etc.). Hence consistency appeared to be related to other factors in much the same way as honesty itself.

There were, of course, many other contributions from this Inquiry. Some of them appear in other sections: cooperation, inhibition, persistence, and reputation. The development of the excellent battery of tests now published by the Association Press is foremost. Evidence on the lack of association between these behaviors and maturation, health, or economic status is a challenge to many theories. The questioning of programs supposed to develop better character has been wholesome. The study as a whole shows, however, the serious limitation of the test-correlational approach to understanding individual character or educational approaches.

Many other reports of conduct tests of honesty may now be found in the literature. Cheating in some form or other has been detected by pre-scoring of self-scored papers, by neglect to report favorable errors, by observation during tests, by comparison of identical errors, etc. in reports by Gundlach (12, 1925), Chambers (28, 1926), Persing (26, 1926), Yepsen (16, 1927), Fenton (33, 1927), Miller (33, 1927), Bird (34, 1928), Brown (34, 1928), Brownell (34, 1928), Bathurst and others (18, 1929), Newcomb and Watson (30, 1930), and Campbell and Koch (30, 1930). There is general agreement among these studies that a minority (25 percent would be a rough average) of the class cheat, that this minority is below the average in intelligence, that special pressure for grades or degrees increases cheating, that what students say on questionnaires cannot be taken as an index of their cheating behavior, and that lectures on honesty are of doubtful value.

May and Hartshorne (28, 1926), in a preliminary study, used a series of graduated situations in which the form and psychological structure of the situation remained constant, while the barrier to cheating was constantly made stronger. At one extreme the subject could cheat by merely erasing

a pencil check. The difficulty in cheating increased up to the other extreme which involved erasing a total phrase written in ink. This study is of importance because it is one of the few in the history of character testing which has been well analyzed to hold constant the pattern, while introducing one variable. The result was, as might have been expected, that the amount of cheating varied directly with the ease of the technic. The consistency of behavior was marked—the pupils who cheated under the more difficult conditions almost invariably cheated under the easier conditions. The method is worthy of much more extensive application.

The overstatement test, dealing sometimes with book titles, sometimes with household abilities, sometimes with school knowledge, was used by Cushing and Ruch (79, 1927) to differentiate delinquent girls; by Huxtable (34, 1928), and Woodrow and Bemmels (33, 1927; 112, 1927) as a test correlating .6 or .7 with general character ratings; by Terman (12, 1925; 30, 1930); and by Lehman and Witty (30, 1930) and Stoke and Lehman (30, 1930) to show the superiority in character of intellectually brilliant children. Maller's Self-Marking Test can be quickly administered and scored, and provides another index of tendency to exaggerate or raise one's score.

Clark (16, 1927) and Tuttle (18, 1929) tried testing religious education curriculums by a few honesty tests given before and after, with disappointing results. Tuttle's method of using groups, each of which *omits* one phase of the full program, is a valuable modification in experimental procedure, for use in situations in which sharp contrasts of program are not feasible. Watson (111, 1928) made the most extensive application of an honesty test to the measurement of results in a program, the test being a form of the May-Hartshorne S-A Test, in which the pupil can exaggerate his own virtues. In this study of boys in summer camps the honesty items were interspersed among items on attitude similarly phrased but not containing the same principle of super-human virtue. Gain during a two weeks period of Y. M. C. A. camp was seven times its standard deviation. Correlation of gain with length of period was .36. Correlations with other measures were all below .4. An item analysis, using the total test as criterion, shows the better and poorer items of this "honest confessions" type.

Zillig (30, 1930) tested a battery of behaviors such as false report of success, taking articles loaned, and boasting about possessions and parents, in 270 German school children. The results are in agreement with Hartshorne and May's but are less extensive.

The largest number of studies in the area of honesty deal with technics for identifying the individual who is giving false testimony. In addition to the psychogalvanic reflex which will be discussed later, one of the principal measures used has been reaction time in association by such investigators as Marston (12, 1920), Goldstein (12, 1923) and Crosland (107, 1929). This was criticized by Stumberg (12, 1925) who showed a number of ways in which sophisticated subjects may "beat" the test, and by English

(28, 1926) who showed that the two indices of increased time for response and increased variability of response do not agree. Nevertheless Crosland was able, using a variety of signs, to identify delinquents in 90 percent of the cases. All of the other emotional indices are applicable to this problem. Changes in heart rate and in breathing curve were studied by Marston (12, 1917), Benussi (12, 1914), Burt (12, 1921), Larson (12, 1921-22), Landis and Wiley (28, 1926), Larson (16, 1927), Adler and Larson (106, 1928), and Chappell (18, 1929). Landis (109, 1927) gave a summary of the results. A very clever test was devised at Moscow by Luria based on the principle of conditioning. A hand movement was associated with the verbal response in a long series of free-association reactions. Then "significant" words, e. g., relating to the crime, were introduced. The subject might repress the verbal expression and substitute an innocuous response, but the curve of hand movement showed the start, the block, and the re-formulated expression. A combination of heart rate, blood-pressure, breathing curve, inspiration-expiration ratio, lengthened reaction time, muscular movement records in grip and perhaps in speech muscles, psychogalvanic reflex, peculiarity of word association, and the like would probably give an almost certain identification of the persons emotionally disturbed by the stimulus words. It would, however, give no indication of the reason for the emotional disturbance, and would not be indicative in the case of those sufficiently hardened to lie without an inner tumult.

#### Test Materials Now Available

<i>Title</i>	<i>Publisher</i>
Athletic Contest, Series H. (C.E.I.)	Association Press
Attitudes S-A. (C.E.I.)	Association Press
Coordination Test (Peeking) (C.E.I.)	Association Press
Maller's Objective Test of Honesty	Columbia University
Maller's Self-Scoring Test of Sports and Hobbies	Association Press
Puzzles Test, Series H. (C.E.I.)	Association Press
Self-Scoring Intelligence and Achievement Tests: (C.E.I.) Arithmetic, Completion, Information, Spelling, Word Knowledge	Association Press
Self-Scoring Speed Tests (C.E.I.)	Association Press
Stunt Parties, Test-H. (C.E.I.)	Association Press
Tuttle's Honesty Test (Spelling)	Stoelting

#### **Humor**

The methods proposed for measuring sense of humor include diary records; jokes to be arranged in order of funniness, both methods being used by Kambouropoulou (115, 1926) and by Barry (34, 1928); observation of laughter by Goodenough (34, 1928); and the arrangement of pictures, e. g., the Healy Picture Completion (29) in a form to be as funny as possible. The field is reviewed by Diserens and Bonifield (114, 1930).

## Test Materials Now Available

<i>Title</i>	<i>Publisher</i>
Almack Humor Test	Stanford University Press

### Inhibition, Caution, Self-Control

Crane's ingenious "guillotine" (116, 1923) tested the ability of an individual to hold steady despite the appearance that a large weight would fall on his hands. Snow (117, 1926) and Wechsler (15, 1926) used a somewhat similar idea in testing the ability of potential chauffeurs to adjust quickly when the lights went off and the apparatus with a bang "blew a fuse." Laird (12, 1923) tested the ability of students to resist distractions due to razzing. Brown (12, 1923-24) suggested that the ratio of items attempted to items correct on a difficult test might be a measure of caution.

Hartshorne and May (71, 1929) experimented with the ability to inhibit a desire to take candy or nuts from box on desk before the appointed time; ability to keep a "wooden Indian" face while back of neck is being tickled with a feather; ability to look through a scrapbook of fifty funny pictures without smiling; ability to retain immobile face during bad odors, bad tastes, or while looking closely at a rasping spark-showering apparatus; and the ability to stand pain as per the Whipple pain balance. The average intercorrelation was .23; the average correlation with ratings, practically zero. Another series of inhibition tests took place at a party and included avoidance of premature starts on races, on the game called "Crows and Cranes" when children must wait for the full name to be called to know which way to run; disregard of misleading movement suggestions in the game "Simon Says, 'Arms Up'"; ability to keep a straight face in a group while funny stories are being read; discriminatory reaction to a whistle but not to other sounds; ability to carry a gift "snapper" through the hall without snapping it. The average intercorrelation in this group was .03, and there was evidence of rapid adjustment to the situation if the test were repeated. The tests most widely used included ability to stop reading an interesting story without breaking the seal which just at the moment of greatest suspense interferes with further progress; ability to refrain from touching a toy combination-safe set on the desk for later use; ability to refrain from touching an attractive group of puzzles, carefully arranged and placed on the desk for later use; ability to carry on simple arithmetic at a rate unaffected by the fact that the margin of the sheet also contained drawings, sensational news headlines, comic strips, riddles, etc. It is not clear just in how far some of these are important character attributes. Reliability of one interrupted story test was .48, four safe-tests correlated .50 with the puzzle manipulation, two picture inhibition tests, .42 with one another. The average intercorrelation of pictures, puzzles, safes, and stories was about .20. Reliability of the composite was .80; its agreement



with reputation, .40. Application of the tetrad difference criterion (183, 1927) gave evidence of a general factor in the several individual tests of inhibition. Total inhibition score showed no correlation as high as .30 with any other measured factor, but the tendency of classrooms to vary as units was shown by a correlation of about .4 between average boy and average girl in the same classroom. This may be explained largely by the influence upon one child of the way in which he sees the others, perhaps the leaders, behaving. Correlation with general all-around character proved to be .38. The considerable attention given to inhibition is probably an outgrowth of Roback's convincing argumentation in favor of "inhibition of impulses in accord with a rational principle" as the essence of character. It is to be noted that relatively few of the tests of inhibition follow that definition all the way through.

Washburne's (118, 1929) suggestion of a test in which the subject chooses one chocolate bar now or several next week, etc., is based on a slightly different principle, that of inhibiting immediate impulses for the sake of larger but more remote gains. Washburne found the test to differentiate significantly between delinquents and non-delinquents, especially at lower mental ages.

#### **Test Materials Now Available**

<i>Title</i>	<i>Publisher</i>
Ruggles Distraction Test (C.E.I.)	Association Press
Speed Test, Series I (C.E.I.)	Association Press
Stunt Parties Test, Series I (C.E.I.)	Association Press
Stories, Puzzles, and Safes Test (C.E.I.)	Association Press

#### **Interests**

Intimate knowledge of an individual's interests corresponds reasonably well to knowledge of the personality. The doctrine of interest in education, the competition among character building agencies for free time, the freedom of vocational choice, have all augmented the attention which needs to be given to the analysis of individual interests. The most obvious technic, that of asking children what they like to do and offering a checklist of suggestions, was followed in the field of recreation by Lehman and Witty (123, 1927) and by a number of religious educational agencies (36; 37). The largest study is one carried through by the Board of Young People's Work of the Methodist Episcopal Church. They combined checklists, standardized discussion procedures, and interviews with leaders of youth. Other questionnaires on recreational interests have been developed or used by Guillet (12, 1907), Kuper (12, 1912), Hall (12, 1907), Poull (12, 1922), Pruette (12, 1924), Courtenay (33, 1927), Jerrel (33, 1927), Trow (34, 1928), and Sonquist (37, 1931). Sonquist's study is particularly noteworthy because he demonstrated that more effective program

counselling can be given to young men by means of a test like his than would be possible for those same leaders without such a device. Interest in the movies was studied by many, among whom are Miller and Abbott (33, 1927). Reading preferences were collected by Jordan (121, 1926), Terman and Lima (127, 1926), Bell and Sweet (12, 1916), Dunn (12, 1921), Wheeler (12, 1920), Kimball (28, 1926), Severance (28, 1926), Huber (33, 1927) and doubtless by many others. Waples (30, 1930) compared questionnaires with actual reading practice, considerably to the discredit of the former.

Questionnaires on preferred subjects at school and preferred occupations have been so numerous and can be so simply reproduced that it is not necessary to list the studies. The problem of the permanence of such choices has been investigated by Crathorne (12, 1920), Franklin (12, 1924; 28, 1926), Willet (12, 1919), Mackaye (16, 1927), and Strong and McKenzie (18, 1930). In every case the reliability seems surprisingly high. None of the studies covers, however, a span of more than a few years, and most of them are in the prevocational period.

Tests purporting to differentiate occupational types have been prepared for occupations in general by Miner (125, 1926), and in a succession from Freyd (12, 1922) to Cowdery (28, 1926) to Strong (15; 16; 17; 28; 30; 33; 34; 126, 1927). The Miner Test offers paired comparisons: work indoors versus outdoors, work that requires planning or work that is laid out, work alone or with others, etc. Strong's blank includes choices from lists of occupations, amusements, school subjects, types of people, etc. Reliability is .85, even after more than a year's time (21). It is scored on the basis of the actual interests of men in the twenty or more occupations for which keys have now been developed. A typical bit of evidence, to be duplicated many times in the articles by Strong, is that the test differentiates certified public accountants from lawyers so clearly that only about 6 percent of either group can rate A for interest in the other. Among thirty-six men rated by three agencies the test was 100 percent right in identifying the failures and 73 percent right in identifying the successes. In a group of life insurance salesmen 40 percent of the A rating men were selling more than \$200,000 a year, while only 8 percent of the B grade men did so well.

Several tests try to differentiate general types rather than specific occupations. Freyd (12, 1922-24) studied some of the differences between the socially and mechanically minded. Murphy (12, 1917) even earlier worked on association differences between literary and scientific minded persons. Wyman's Interest Test (129, 1929) differentiates intellectual, social, and activity interests with a reliability over ten days of about .7; over five years Terman found that it dropped to .3. Garretson (30, 1930) developed an excellent blank for differentiating the high-school pupils with a bent toward technical training from those who incline toward commercial or academic study. He and Symonds make the point that this test in no

degree connotes ability in the chosen field, but only line of preference. The relative standing in achievement of persons whose preference leads in one direction will be determined by intelligence and other capacity measures.

In still more specialized form the Minnesota Mechanical Abilities Tests include from Freyd and Hubbard a blank on interest along mechanical lines (30, 1930); the Morris Trait Index L attempts to differentiate by interests the students likely to show leadership in such a profession as teaching; Hendry (37, 1930) and his associates have developed an interest blank which is related to success in boys' camp leadership.

One good way to discover an individual's interests is to find out how he spends his time. Time schedule studies were made by Forman (23, 1926), Martin (33, 1927), Newcomer (33, 1927), Sturtevant and Strang (33, 1927), Bridges (18, 1929), Coy (30, 1930), and Andrews (30, 1930). Actual behavior observation was used by Augustin (33, 1927), Eckstein (33, 1927), Bridges (16, 1927), Hulson (30, 1930), and Ehrle (30, 1930); the subjects were largely young children in a play environment. What people talk about as an indication of their interests was recorded by Landis (122, 1927) and by Stoke and West (30, 1930).

Granted reasonable equality of opportunity, the items of knowledge acquired by an individual give something of a clue to the direction of his interests. One of the most ingenious and valuable interest tests, now, unfortunately, quite outmoded, is Ream's test of range of interests, based upon the extent to which an individual was familiar with terms used in church, in poker, in golf, chemistry, boxing, finance, and such other fields. Pressey utilized some of this idea in her Sports Information Test (21) not yet standardized. McHale (30, 1930), similarly, used the items of information picked up by college girls as an index of their vocational bent.

Other approaches have also been used. Records of the extracurricular participation of the individual are often valuable in vocational guidance, and form the basis of studies by Stanforth (16, 1927), Terry (16, 1927), Kauf (33, 1927), Thornhill and Landis (34, 1928), and Chapin (232, 1926). Choice of companions described in terms of certain characteristics, thus giving a clue to interests, forms the basis for tests by Raubenheimer (12, 1925), Cushing and Ruch (79, 1927), and Watson (111, 1928). Burt (12, 1923) used among other tests the crossing out of irrelevant words interjected in prose which dealt with various topics, the theory being that the interest in the topic would lead the reader to skip more of the words which should have been crossed out.

No question in the field of interest has been more disputed in scientific studies than that of the relationship between interest and ability. Studies have appeared by Thorndike (12, 1912, 1917, 1921), Kitson (12, 1916), Fryer (12, 1923-25), Hartman and Dashiell (12, 1919), Bridges and

Dollinger (12, 1920), Kornhauser (16, 1927), Commins and Shenks (33, 1927), and Wilson (33, 1927) and have been well summarized by Uhrbrock (128, 1926). It is undoubtedly true that in our present social situation dull pupils often look forward to occupations which will bring them a higher social and financial status than they are equipped to fulfill. On the other hand, when these social factors do not play so important a role, as for example, in preferring one school subject above another, or in choosing general types of activity such as work with people, with things, or with ideas, there is good evidence for the correspondence of interest and ability.

#### Test Materials Now Available

<i>Title</i>	<i>Publisher</i>
Find Yourself Blanks	Association Press
"How Do You Feel About It?" Interest Analysis	Association Press
Jones Personnel Questionnaire	Stoelting
Lehman Play Quiz	Association Press
Miner, Analysis of Work Interest Blank	Stoelting
Minnesota (Hubbard-Freyd) Mechanical Interest	Stoelting
Morris Trait Index L	Public School Publishing Co.
Pressey, Sports Information Test	Stoelting
Sonquist Interest Finder	Association Press
Stanford Educational Aptitudes Test	Stanford University Press
Strong Vocational Interest Test	Stanford University Press
Symonds-Garretson Vocational Questionnaire	Columbia University
V.I.Q. Booklets (Hepner)	Stoelting
Wyman Interest Test	Stanford University Press

#### Introversion, Extroversion

The concepts of introversion and extroversion introduced by Jung, have led to a number of attempted measures. As Hendrick (3, 1928) pointed out, most of the measures deal with static traits, whereas Jung had in mind a variable complex mechanism which might show itself in very diverse behaviors. Freyd began in 1924 (130, 1924) with an analysis of the types of fifty-four traits. Heidebreder (28, 1926) analyzed the responses of people to these items and found thirty-one which were apparently consistent and from these built her scale. Other paper and pencil self-report measures were developed from much the same definition by Laird (the Colgate scales, one for self-rating, the other for rating by an observer), Conklin (33, 1927), Whitman (18, 1929), Neyman-Kohlstedt (18, 1929), and MacNitt (21). Marston (133, 1925) was the first to formulate a behavior test to differentiate these types. Attempts to find a difference between introverts and extroverts on the ordinary laboratory psycho-technical measures were made by Schwegler (33, 1927) and Washburn (30, 1930), the former finding introverts somewhat slower in movement

and less rich in emotional output, the latter finding no difference in reaction time, flicker sensitiveness, or extremes of liking or disliking of colors. Hovey (18, 1929) found no consistent differences between the types in performance under distraction.

Newcomb (134, 1929) collected several thousand situation observations on fifty-one boys in a summer camp, under the guidance of leaders given special psychological training. The tendency for a characteristic behavior noted on one day to recur in what were apparently similar situations later, was very slight ( $r = .2$ ). The tendency for specific behaviors to be consistent with the usual trait names, and for the traits to cluster in such aggregations as introvert-extrovert types was similarly weak. He found that when the boys were rated on traits by the leaders, the coherence of the traits and types was very much more evident than it was by sorting the concrete observations. He concluded that the apparent consistency was imposed upon the facts by the expectations of the leaders rather than an outgrowth of the behavior itself. Another possible explanation would be that the leaders saw a form or structure in the behavior which was, to their minds, consistent and properly called by some trait name, but which could not be made evident in the brief behavioristic explanations. The study is so fundamental for all kinds of character analysis and test-making that it deserves repetition and variation to establish the real nature of the situation.

The concepts of introversion and extroversion are rather vague and have been subject to many definitions. Control of temper is characteristic of the extrovert, some would say; of the introvert, according to Hewlett and Lester (34, 1928); of neither particularly, according to most writers. Hence the various tests and indices of introversion agree poorly among themselves, as has been demonstrated by Guthrie (132, 1927), Broom (18, 1929), and Weber and Maijgren (18, 1929).

Measures of something called introversion and extroversion have been used by Caldwell and Wellman (28, 1926) to show that leaders are more likely to be extrovert and by Billingsrath (30, 1930) to show that leadership has a zero correlation with introversion-extroversion. Conklin (33, 1927) found that salesmen are the extreme extroverts, but Gallup (28, 1926) found the test of little value in selecting successful salesmen. One study found the sick more introvert than the well, another found advanced tubercular patients the most extrovert. Heidbreder (33, 1927) and Oliver (3, 1930) agreed in finding no sex differences; others have found women more introverted. Kovarsky (34, 1928) found the older, duller boys more introverted, but Oliver and most others found no differences related to age or intelligence. There is disagreement on the relationship of this dichotomy to handedness among such investigators as Downey (16, 1927), Estabrook and Huntington, and Wetmore and Estabrook (18, 1929); on its value in diagnosing the insane by Campbell (18, 1929) and Neymann and Kohlstedt



(18, 1929); in characterizing nurses by Elwood (16, 1927) and South and Clark (18, 1929); and in yielding a difference related to extent of participation in sports by Steen and Huntington (18, 1929) and Hewlett and Lester (34, 1928). In short, wherever two or more studies have been made, the results appear to be conflicting. Perhaps the difference would disappear on more careful analysis of the sort of introversion being studied, the group studied, and other influential factors.

Further suggestions, appearing each in only one report and hence subject to great tentativeness, are the findings by Conklin, Byrom, and Knips (16, 1927) that extroverts are likely to have less severity in menstrual upset; by Wells (28, 1926) that they are likely to be more promiscuous sexually; by Conklin (33, 1927) that journalists are likely to be introverts and business administrators extroverts; by Davenport (16, 1927) that inspectors are more introverted than foremen; and by Young and Shoemaker (34, 1928) that those who are intelligent and introvert select literary majors, while those who are intelligent and extrovert select chemistry or biology majors. Most studies yield a normal distribution, with introversion and extroversion evident only at the extremes. Downey (28, 1926) found that to be true of the members of the American Psychological Association. Oliver (30, 1930) found no correlation between introversion-extroversion and age, intelligence, ascendance-submission, prejudice, or social intelligence, but some difference in scholastic and emotional traits. Hunter and Brunner (34, 1928) found no relation to gambling. Hewlett and Lester (34, 1928) found introverts (rather unusually defined as those who regard themselves as self-controlled, worrying, lacking grit and initiative, etc.) to be lower in I. Q., poor in health, but not different in recreations, or position in family. Sonquist (37, 1930) built a criterion for the success of young men in leading groups of boys in Chicago, but found introversion-extroversion measures unrelated to it. Similarly Hendry and others (37, 1930) found that successful camp leaders could not be regarded as markedly introvert or extrovert in trend.

The many shades of variation in theory as to what these two terms, *introversion* and *extroversion*, really should mean need not concern us here. They have been well summarized in Roback's *Psychology of Character* (277, 1927) and in the review by Guilford and Braly (131, 1930). The main contributors have been Stern, Klages, Jung, Hinkle, White, Wells, McDougall, Downey, Conklin, Kempf, and Hunt; each has modified the theory a little to emphasize some special interest or category. There is no evidence on the innate character of the traits, but about fifty-fifty division among the theorists in their assumption of the importance of heredity. Knight Dunlap a few years ago made a vigorous attack upon the whole concept and urged that this illegitimate offspring of psychoanalysis, which had been laid on the psychologists' doorstep, should be placed in some sheltered institution for the unfit.

### Test Materials Now Available

<i>Title</i>	<i>Publisher</i>
Bernreuter Personality Record	Stanford University Press
Colgate Mental Hygiene Scales C-2, C-3	Hamilton Republican
Heidbreder-Freyd Introversiion Scale	Heidbreder
MacNitt Psychological Interview	MacNitt
Neyman-Kohlstedt Diagnostic Test of Introversion-Extroversion	Stoelting

### Leadership

Leadership tests are found best in the actual life situation. Moxcey (137, 1922) used financial success in the ministry, Terman and Cox (138, 1925-30) used recognition in encyclopedias, many have used inclusion in *Who's Who*, Bowden (135, 1926) and others used election to student council chairmanship, Goodenough (136, 1928) used observation of children in free-play periods as indices of leadership. The consistency of the short-sample series of behavior observation in the last named study was better than .8. Sonquist (37, 1930) used ratings on leaders of boys' clubs; Hendry (37, 1930) used ratings on leaders in summer camps; Morris used success in practice teaching as a criterion for her test (18, 1929). Most of these studies attempted to find some other indices which could be used to predict success in leadership. Among these are intelligence, relative youth, extroversion, larger size, superior scholarship, superior behavior, more sociable and talkative behavior, more extra-curricular participation, liberal or no religious affiliation, attractive appearance, higher ratings on industry and ambition, more interest in the aesthetic, more contact with modern social issues, etc. No studies have so far proceeded from the selection of characteristics of leaders in one situation to see how well, by means of those characteristics, leadership in a slightly different group could be predicted.

### Test Materials Now Available

<i>Title</i>	<i>Publisher</i>
Hendry Camp Leadership Test	Association Press
Morris Trait Index L	Public School Publishing Co.

### Maturity: Social and Emotional

One phase of social maturation is the ability to interpret correctly shades of emotional expression in others. G. S. Gates (141, 1923) tested this first with pictures and later (12, 1925) with vocal expression from a phonograph record recitation of the alphabet. Dashiell (33, 1927) suggested an improvement in the technic by the use of stories to convey the emotion which is to be matched with voice or appearance. The stories are less dependent on the development of abstract ethical or psychological vocabulary. A. I. Gates (140, 1924) checked another theory, namely, that social and emotional maturity could be measured in terms of mental age or in terms

of such physiological indices as the ossification of the wrist bones. Inter-correlations with ratings were so low as to suggest that one form of maturation cannot be used as an index of any other. Gessel and Lord (33, 1927) measured a type of maturity by observing the child's ability to care for his own person, button his clothes, etc. Children from higher social levels were superior in spontaneity and responsiveness; children from lower social levels apparently developed independence and technic of self-care at earlier ages. Chambers (12, 1925) modified the Pressey X-O norms by finding items characteristically liked or disliked at each age level. It was then possible to score the child's response in terms of his emotional maturity. Furfey (139, 1928) for several years has been working on measures of the factor which differentiates two twelve year old boys of like intelligence, but one of whom seems still babyish, the other very mature in his attitudes and behavior. He used a rating scale of eighteen items which, combining all items and the estimates of two judges, gave a reliability of .94. The test included lists of imaginary books, of play activities, opinion records, etc. and had a reliability of .76. The correlation of the test with ratings was .56, but with M.A. was only .23. Weber's similar measure of emotional age correlated .4 with M.A. and .5 with C.A. (30, 1930).

#### Test Materials Now Available

<i>Title</i>	<i>Publisher</i>
Furfey's Child Development Test	Stoelting
Gates Test of Social Perception, inclusive records	Stoelting

#### Moral Knowledge, Ethical Judgment

The earliest development in character testing (aside from reputation measures) and the one which comes first to mind when character tests are mentioned, is the investigation of what the individual thinks about matters of right and wrong. Before the beginning of the Twentieth Century professors of moral philosophy occasionally gave an inductive turn to their studies of moral ideas (153, 1898; compare also Brogan, 12, 1923-25). Fernald's list of offenses (145, 1912) to be arranged in order of seriousness, appeared in 1912; and the same idea has since been used by Bronner (12, 1914); Tanaka (28, 1926); Weber (28, 1926), who found no great difference between female delinquents and Wells College girls; Snyder and Dunlap (12, 1924), who increased the list to one hundred acts to be rated; Quadfasel (15, 1925); Pitkin (15, 1926), who used the Ten Commandments as material to be ranked; Slavens and Brogan (33, 1927), who obtained rankings both on frequency and badness; Rosner (16, 1927), who had the acts to be ranked described on cards to facilitate sorting; Thurstone (33, 1927), who built a scale of equal units based on the rankings; and Palluch (34, 1928), who used good and bad acts in lists and also in stories.

The most common type of test for moral judgment presents a situation and asks for the subject's choice of the best response. Some of the more naive tests have asked what the individual tested would do in such a situation. As a rule, however, test makers have been aware that good answers might arise from good practice or conscious or unconscious pretense, while poor answers might arise from lack of knowledge of the right, or from a sense of honest humility about one's achievements. This ambiguity is wholly undesirable and has led to the general acceptance of the type of question which asks the subject what he believes it best to do. It does not ask whether he does it or not. The object is not to find out how the individual behaves, but rather his familiarity with the approved standards. Tests of this type were used by Kohs (21; 147, 1922); by Athearn (142, 1924) in testing the Sunday School pupils of Indiana; by Watson (154, 1926; 111, 1928) in testing about 15,000 boys in Y. M. C. A. groups; by Descoeudres (2, 1914); by Patrick (15, 1926) in studying race differences; by Johnston (15, 1925), who compared snap judgments with more carefully reasoned choices; by Hoyland (15, 1926) who tested a thousand children in India; by Blomfield (16, 1927), who tested church school pupils; by Katz (34, 1928), who studied ideas of cribbing and other college practices; by Jones (18, 1929) in his study of disagreements; by Boynton (18, 1929); by Tuttle (18, 1929), who studied the contribution of religious education; and most extensively by Hartshorne and May (146, 1930).

Before taking up the contributions of these and related studies it may be well to complete the list of types of test. Woodrow (155, 1926) used pictures representing children in acts of service or destruction, four to the page, and asked children to choose the picture they liked best. The reliability of eleven such pages was .79; correlation with ratings on general character in primary grade children was .41. Chassell (144, 1924) experimented with a test in which pupils were asked to weigh the consequences anticipated from each of several lines of action. The theory underlying such a test is, of course, that the best character is the one who is able correctly to see and to evaluate the consequences of his acts. Patterson (146, 1930) has given the best elaboration of this approach in a series of tests to measure foresight of consequences. The foresight tests showed a correlation of .6 with intelligence; with other moral knowledge tests .5, with school marks .4, with honest conduct .4, with emotional stability .3, with persistent, helpful, or controlled behavior, correlations of about .2. Correlations with general all-around character was .50. Schwesinger (152, 1926) studied ethical vocabulary, finding that understanding of the terms commonly used in describing and analyzing right and wrong behavior gave a correlation of .9 with intelligence and of practically zero with honesty. Watson (111, 1928) used this fact to include a test of intelligence (ethical vocabulary) within a test that looked to be entirely a morality test and so secured a measure of intelligence in a situation in which intelligence tests would not have been welcome. Eastman (28, 1926) studied the information about current social

life possessed by delinquents and found that they compared very favorably with non-delinquents of similar mental age. Williams (16, 1927) asked junior high-school pupils to list their heroes, and found their ideal personalities limited to the conventional types from history books and Sunday School, with a few modern actresses or ball players.

A considerable number of other studies have been more interested in qualitative analysis of the moral attitudes of individuals than in quantitative scores. Among them may be mentioned McGrath (12, 1923), Mitchell (12, 1925), Sharp (12, 1908), Tudor-Hart's study of cases in which lies are believed to be necessary (15, 1926), Macaulay and Watkins (150, 1926), and Studencki (15, 1926), who studied children's ideas of what made one good and what made one wicked. Dearborn (33, 1927) questioned ideas of honesty, and Sanaryahz, working with Belsky (33, 1926), tested the reaction of children to realistic situation-descriptions, aiming at insight into social values rather than at a score.

Brotemarkle's (143, 1922) technic consists in arranging a series of words describing degrees of a trait, e.g., between bravery and cowardice, in a rank order. The net result is, of course, some measure of the similarity or difference between the way in which the subject tested interprets shades of meaning in these terms and the way in which the words are arranged in Brotemarkle's norms. That it has any other significance is doubtful.

The use of knowledge measures has not been carried far into knowledge of the psychological and social conditions influencing behavior. May (12, 1920) used knowledge of the underlying Scripture and philosophy to differentiate conscientious objectors who had the sort of training which they said led them to their ideas, from those who assumed this position as a way of evading military service. In the moral knowledge tests of the Character Education Inquiry (146, 1930) are sections on understanding of some common cause and effect relations, on ability to identify acts properly called cheating, lying, or stealing, as well as the vocabulary and foresight tests mentioned above. Agreement or disagreement with some principles of conduct and some generalizations about others are included.

Among the types of opinion studied (146, 1930) are opinion as to what is one's duty, what is the best thing to do in a given predicament, whether acts usually condemned might be justified under certain provocations, which consequences are most probable and which most serious, choice of conflicting values, e.g., immediate versus remote, personal versus social, physical versus spiritual gratification, etc. Liao (12, 1919) and Chapman (12, 1920) studied not the solution which pupils would give, but their reasons or rationalizations for the solution given by the author as good.

Lincoln and Shields (148, 1931), following the Binet principle, constructed an age scale for measuring moral knowledge. It may be seriously doubted whether morality improves with age (154, 1926; 146, 1930), but certainly the scale is right in suggesting that the situations in which it is to be measured change with age. The scoring in terms of a moral knowl-



edge age would be most unfortunate. The standards for this scale, as for many of the others, appear rather arbitrary and subject to considerable dispute.

One of the first questions arising with reference to tests of the opinions and judgment of children is that of sincerity. Naturally the conditions of the test make some difference at this point. In general, however, when children are asked what they think it best to do under given circumstances, or how they think certain actions should be rated, they give a genuine report of their notion of what is expected of them. Clearly it is not a report of what they do. It must never be so interpreted. It is what they think the moral obligation is. Hence it is not surprising that two-thirds of the answers given by children in disagreement with what they were later told was the "right" answer, were maintained by the child in spite of his knowledge that the code said something else. He still thought he was right and stood by it. This obviously is something more than a mere desire to please the examiner. Hartshorne and May used the correlation between moral knowledge scores and scores made on the S-A Lying Test in which a pupil does try to exaggerate his virtues, and found no relationship ( $r = -.05$ ). Obviously high moral knowledge scores are not made by the kind of pupil who pretends to be very virtuous.

The next question, "What is the relation between moral knowledge and actual conduct?" has not been satisfactorily answered. Chambers (28, 1926), and Brown and Shelmadine (34, 1928), found that pupils might agree on condemnation of cheating and still cheat. Sorokin (34, 1928) and Stabler (18, 1929) found only partial accord between social ideals and conduct. Yet Katz (34, 1928) found that those who cheated believed that others did so, while those who did not tended to believe others also honest, an application of the principle of projection which has been used in some moral knowledge tests. Between moral knowledge and general all-around character Woodrow (155, 1926) and Watson (154, 1926) found that there is a definite positive correlation. This is nicely confirmed by the fact that general all-around character as measured in terms of all of the tests and scales given by the Character Education Inquiry showed a correlation (.6) with moral knowledge higher than the correlation with the various conduct tests, and definitely higher than the correlation with intelligence or home background measures.

There are two approaches to the further study of the problem. One is comparison of a situation described in such a way that it is a very exact psychological and dynamic parallel to the situation experienced. That has rarely been found. Persing (28, 1926) found that of 87 percent who said they would report papers which they found to be graded too high, only 21 percent did so. On the other hand (146, 1930) "ten of eleven answering 'Let another pupil copy your work and say nothing about it,' did actually cheat on tests. All of the five who answered that the best thing to do with

an obdurate slot machine was to smash it and get your nickel back, did actually cheat. Thirteen of the fourteen who approved John's cheating on a test to help his class win, did cheat on the tests." But these same items repeated on a new population showed a differentiating power of only three standard deviations between honest and dishonest (still a large difference) as compared with ten standard deviations on the original group. So far as these results go they suggest the hypothesis: A good answer on a moral knowledge test is not evidence of correspondingly good conduct in the actual situation; a poor answer on a moral knowledge test is strong indication of a poor response in the conduct situation.

The other approach is statistical, and was the main one used by Harts-horne and May (146, 1930). The difficulties of interpreting such results are suggested by their summary:

The relation between all moral knowledge tests and all conduct tests may be said to be .12 or .35 or .63 or .87. The first is limited to one population (Y) and is obtained by using scores as deviations from classroom means. The second is also based on population Y, but uses as scores, deviations from the mean of that population. The third is like the second except that it is based on all three populations, thus raising the correlation from .35 to .63. Using classroom means in population Y as units, the correlation would have been .84.

Clearly a correlation of .12 would argue one way, a correlation of .84 the other; and both are true. The relationship in these data is dependent upon something which goes by classroom and population groups rather than on a tie within the individual. Good codes go with good conduct when groups are taken as a whole; that does not follow in the comparison of individuals from the same social and school class.

All of the studies are in agreement in showing a considerable relationship between moral knowledge and intelligence, usually from .4 to .7. The findings of Athearn (142, 1924), Watson (154, 1926), Hartshorne and others (146, 1930), Franklin (34, 1928), Shuttleworth (33, 1927), Blomfield (15, 1927), Tuttle (18, 1929), Moran (18, 1929) and best of all, Hightower (30, 1930) are in more surprising unanimity in showing that there is no close relationship between Biblical or other religious training and moral judgment tests, provided intelligence of pupils is kept reasonably constant. Moral knowledge tests seldom are useful in differentiating delinquents from other subjects of like intelligence and environment as shown in studies by Bronner (12, 1914, 1922), Lowe and Shimberg (12, 1925; 149, 1925), Weber (28, 1926), and Palluch (34, 1928). Attempts to test race and national morality have shown some slight differences, Patrick (15, 1926) and Hoyland (28, 1926) especially suggesting that material considerations are less evident in India than in America. No generalizations would so far be warranted.

In two directions at least, progress is being made in defining the concepts used in the moral knowledge test. It is evident enough from such studies as that of Jones (18, 1929) that people disagree more on some moral

questions than on others. Maller made a fairly extensive survey in one part of one culture of the amount of agreement. He proposed to reject questions (for test purposes) on which everyone agrees, to use for discussion material in moral training questions on which adults and children differ markedly within their own age groups; and to use for tests, questions upon which competent adults agree and children show a range of opinions. The other direction of promise is Carmichael's collection (18, 1929) of the moral problems confronting six-year-old children which may make possible the construction of a test with a minimum of artificiality in its contents.

Several experiments in the use of moral knowledge tests in clinics point to a value in discussing with the subject *why* he answered as he did, that is much greater than the value of the numerical score. It may be expected that the years ahead will witness an increasing use of such tests as instruments of individual re-education and guidance.

#### Test Materials Now Available

<i>Title</i>	<i>Publisher</i>
Baker, Telling What I Do Test	Public School Publishing Co.
Brotmarkle, Moral Concept Test	Stoelting
Good Citizenship Test	Association Press
Hill Test of Civic Attitudes	Public School Publishing Co.
Information Tests, Forms I and II	Association Press
Kohs Test of Ethical Discrimination	Stoelting
Lincoln and Shields, Age Scale	Shields
Opinion Ballot A, Forms I and II	Association Press
Opinion Ballot B, Forms I and II	Association Press
Wilson Ethical Discrimination Test	Stoelting

#### Morphology, Constitutional Type, Physical Build

Kretschmer's original suggestion (157, 1922) that individuals of asthenic build tend toward the schizoid in personality and that individuals of pyknic build tend toward the cycloid (manic-depressive) temperament has been fairly well confirmed in later studies by Weil (12, 1922), Gurewitsch (3, 1926), Yezlin (3, 1926), Polen (3, 1928), Mohr and Gundlach (156, 1927; 15, 1929), and Wertheimer and Hesketh (15, 1927). Some of these, notably the last named, varied the index somewhat. Adler and Mohr (3, 1928) were critical, feeling that only extremes can be so classified into types. Kroh found his results with schizothymes in direct contradiction to Kretschmer. Naccarati in his own studies and those with Garrett (158, 1923; 159, 1924; 12, 1924) found type unrelated to intelligence tests or to ratings on character in normal persons. Sheldon (162, 1927) similarly found no evidence that trait ratings correspond to physical measurements of any sort. Bender (34, 1928) disagreed with the conclusion of

Kitson (12, 1922) and Snow (28, 1926) that successful salesmen are above the average in weight and height. Graves (15, 1926) suggested a new type classification based on scapular formation. Paterson (160, 1930) in 1930 gave a good critical review of the evidence.

### Opinions, Attitudes, and Prejudices

As more attention is given to the society in which men move as the source and result of individual character, larger and larger importance must be given to those attitudes which represent the contribution of the individual to this social setting. It is no longer enough to be a good character in face-to-face relations; there is a demand for character extensive enough to include economic class, national, international, and race relationships. The importance of this phase of character, considered in conjunction with the relative ease of test-building for opinion study, has led to a very large number of contributions. Sumner in 1898 published a report of investigation of beliefs which was probably not the first to be made. Woolston (12, 1916) early made use of the method of arranging names of nations in the order of preference. Young (33, 1927) and Thurstone (34, 1928) were responsible for later developments of this same technic. Thurstone presented a technic for scaling the responses so that units might be equivalent. The Bogardus Social Distance Test (165, 1925) is an ingenious modification of the rank-order approach. In this test the subject indicates the degree of his antagonism to the nationality or race or class or religious group named, by checking along a scale of intimacies ranging from admitting such persons to the country to admitting them to the family by marriage. Further work, using the Bogardus scale, was done by Park (37, 1925); Binneweis (15, 1926), who studied rural groups; Poole (15, 1926), who studied personal versus social groups as stimuli; Wilkinson (18, 1929), who studied occupational groups; and Woolston (34, 1928), who found that a variety of tolerances were more apt to exist in the younger, politically non-partisan, liberal arts, non-church-belonging students.

The most common technic has been the collection of a series of statements to which the subject may respond by agreeing or disagreeing. Sometimes degrees of accord are provided. Not always are the scales as carefully constructed as was the Allport-Hartman study (163, 1925). These authors studied student essays on the various topics to be used and selected statements in the words of the students themselves which could then be roughly graded from the most extreme radical to the most extreme reactionary opinion. These statements formed the scale, one made from actual samples and not merely from the scale-makers' own imagination. Other more or less inclusive tests of opinion have been tried out in the studies of Folsom (12, 1919), Wadmore (12, 1922), Hart (12, 1923), Symonds (12, 1925), Jones (28, 1926), Lundberg (26, 1929), Reed (16, 1927),

Moore (172, 1925), Washburn (16, 1927), Vetter (175, 1930), Harper (169, 1927), Willoughby (30, 1930) and, presumably, many others.

Most of the statement scales have been produced in certain limited areas of attitude. International attitudes, for example, were tested by Keeny (28, 1926) with a Bogardus test modification; by Frederick (33, 1927); by Diggins (33, 1927); by Neumann (28, 1926); by Watson and others in the Y. M. C. A. Test of Opinions on International Questions (21); and in the Orient and Occident study (176, 1927); in several of the attitude scales being developed at the University of Chicago by the Thurstone method; and most exhaustively by Heber Harper (168, 1931) in his study of the attitudes of students in Europe and America. Religious attitudes were studied in statement scales developed by Case, Bain (33, 1927), Shuttleworth (33, 1927), Sturges (33, 1927), Betts (34, 1928), Ford and Starbuck (37, 1929), Starbuck and Sinclair (33, 1927), Howells (30, 1930), Van Ormer (33, 1927), and still more specifically in MacLean's study of children's ideas of God (171, 1930). Acceptance or rejection of superstition was the center of attention in studies by Garrett and Fisher (28, 1926), Fisher (33, 1927), Miller (18, 1929), and Lundeen and Caldwell (30, 1930), the latter collecting reactions from nine hundred high-school pupils on matters like predicting hard winters from squirrel activities. Among the other areas in which statement tests of attitude or opinion have been utilized may be mentioned race attitudes by Busch (15, 1926), Orata (16, 1927), and the Y. M. C. A. Opinions Test (21); attitudes toward law by Lockhart (30, 1930); moral beliefs by Dudycha (30, 1930); attitudes toward offenses committed by school children by Wickman (164, 1928); sex attitudes by Rice (164, 1929) and Davis (166, 1929); attitudes of young people toward home by Burger (30, 1930); attitude of teachers of education toward military training by Coe (33, 1927); and attitudes toward a variety of matters taken up in conferences (177, 1925; 36, 1929; 37, 1930). Here should be mentioned also Elliott's *The Process of Group Thinking* (167, 1928).

Some studies are based on questionnaires without suggested answers. There is no reason for believing that the suggested answers give a more accurate picture of the subject's ideas than the ordinary questionnaire about which many derogatory remarks have been made in scientific journals. The subject's own expression is likely to be considerably more satisfactory to him than his check on a multiple-choice question. The advantage of the latter, however, is that it requires responses which are comparable from person to person. All that can be maintained on the basis of controlled answer instruments is that those who check the same response are united in preferring it to any of the other suggested statements or degrees of accord. Qualitative aspects of the choice, reasons underlying it, and the like, can be studied in the individual case much better with the free-response questionnaire. Both types of questionnaire are dependent upon the con-



struction of a situation such that the subject wishes to air his real views. Among the opinion questionnaires using free response technic may be mentioned one on patriotism offered by the Bureau Internationale d'Education (35, 1926); a study of attitudes of young business women toward the home and marriage by Cavan (33, 1927); religion by Kupky (34, 1928); international animosity by Baumgarten (34, 1928); the church by Baber and Stroud (36, 1929); religion by Westphal (34, 1928); choice of heroes by Hewlett (12, 1918) and Moore (12, 1920); and of what girls tell their mothers, Leonard (170, 1930). Hamilton's study (101, 1929) of experiences and attitudes of married men and women toward their marriage is a model in the controlled interview technic. Rather surprising is the evidence obtained by Pointer (34, 1928) that more frankness is obtained from married women by an anonymous questionnaire than in an interview with a woman psychiatrist, cooperation being voluntary in both cases.

Attempt has been made to avoid the difficulty of securing complete voluntary self-revelation by creating measures which would reveal the subject's attitudes without his being aware of the revelation. Thus Shuttleworth (12, 1924) showed how the Hart questionnaire of attitudes and interests could be scored to show the "money-mindedness" of college men. The Watson Test of Fair Mindedness (177, 1925) purports to be a test of public opinion, but is scored to show the kind of prejudice which is manifested by emotional reaction to words, by believing doubtful theses so true that no one of sound mind could question them, by drawing emotionally desired conclusions from evidence which is really quite ambiguous, by approving acts if done by one group and disapproving similar acts in another less favored group, by regarding all arguments, strong and weak, on the favored side of the question as strong and all opposing arguments as weak, and by generalizing from a few instances to approve or disapprove a whole class of persons. None of these scoring technics are apparent to ordinary students of psychology in the course of testing. This test has been further used in an unpublished study of newspaper editors in Oregon; in studies of Y. M. C. A. secretaries, e. g., Swift and Pence (36, 1929); of educational students by Clark (37, 1930); and of gifted children by Terman (138, 1925-30). Word association methods were used by Gisp (34, 1928). Extremism was tested by Jones (28, 1926) in a technic not unlike one of Watson's. Weinland (30, 1930) proposed that reactions to proverbs can be used as a test of conformity or variability. Lentz (30, 1930) similarly proposed that an opinion test can be scored to show conservatism, acquiescence, and variability. Reed (16, 1927) used an opinion test in which he was interested primarily in consistency of trend toward radical or conservative replies. Watson (29) described a test dealing with international relations which was scored only to show the number

of paired contrasting statements, scattered through the test, which had been consistently answered. Direction of answer was neglected, but the assumption was that no straight-thinking individual should agree with both of the statements. The best use, and probably the original creation, of the technic of consistency within the test as a score, is found in Manley Harper's study (169, 1927) of the attitudes of educators. He found evidence for believing that the radicals were better educated, were more consistent in their answers, and were more critical, the last being evidenced by a tendency to make fewer of their scores by agreeing with the proposition presented.

Thurstone's contribution (34, 1928; 174, 1929; 18, 1929) has been in the direction of improving the units of measurement and not in the direction of concealing the purpose of the approach. Statements are sorted by judges according to one linear scale from the most extreme in one direction to the most extreme opposite. It has been shown in some of these studies that statements are sorted in practically the same piles by judges who are of one opinion and by judges of quite contrary opinion. Hence the position of statements on the scale is acceptable more or less regardless of the point of view of the one who sorts them. Ambiguous statements, statements that extend out in other dimensions but are not clearly placed with reference to the underlying linear scale of attitude, can be eliminated. As a result of this more accurate scoring of each response, it is possible to obtain reliable indices of attitude in much shorter compass than was true with the older and cruder scales which simply assembled interesting statements. Of course, the correlation between attitudes tested by a series of statements which one might write down offhand, in half an hour, and attitudes tested under the same circumstances by the more refined scale technic would probably be above .9. The advantage of the Thurstone Scales is that, once developed, they take less time to administer and to score, and give a report which is less padded with errors. Scales developed by Thurstone's technic have been (or are being) prepared in areas of attitude toward God, the church, war, Negroes, birth control, movies, Chinese, Germans, the U. S. Constitution, law, Sunday observance, prohibition, censorship, criminals, communism, patriotism, public office, capital punishment, labor unions, economic position of women, divorce, evolution, social position of women, immigration, free trade, German war guilt, preparedness, freedom of speech, the League of Nations, the Monroe Doctrine, and foreign missions.

A common type of study based on opinion measures is correlational. Along with the opinions are collected other data which permit of computation of the extent of relationship between certain attitudes and certain possible conditioning factors. For example, Allport (163, 1925) found that radicals and reactionaries were much alike in being more emotionally unstable and more inclined to overestimate their intelligence than were

the middle-of-the-road group. Lundberg (164, 1927) found that attitudes as measured by votes cast were related to the economic conditions of the area, the more radical vote coming from less prosperous counties. Moore (172, 1925) found little difference between radical and conservative undergraduates in intelligence or emotional stability, but found the radicals more resistant to the pressure of majority opinion, quicker in reaction time, better able to acquire new habits, and more unusual in their word associations. Washburn's repetition with women students did not confirm Moore's conclusions (16, 1927). Little influence of church training on international attitudes appeared in the studies by Keeny and Watson (28, 1926). International attitudes appeared in the study of Diggins (33, 1927) to have close correlation with distribution of friends among other nationals, but little relationship to familiarity with the language or to travel abroad. Orata (33, 1927) found race prejudice reduced among older students, among those with cultural interests and with friends of other races. The most extensive published study of this sort is Watson's survey for the American Group of the Institute of Pacific Relations, reporting on about three thousand Americans in every walk and station of life (176, 1927). It appeared that attitudes of Americans toward Japan and China were predominantly friendly in a proportion about three to one at that time; that the correlation between being a well-informed group and being a liberal group was .8; that geographic location was much less important than is commonly supposed in influencing opinion, while social or economic class plays a very important role; that travel abroad was not especially important, but that friendship with Orientals was a good index of liberal spirit; and that the attitude of a group could be determined with considerable precision from the proportion of its magazine reading which came from a list of liberal as compared with a list of conservative publications. Lundberg (28, 1926) made the stimulating observation that opinion of citizens did not agree as closely as is popularly supposed with the editorial viewpoint of the newspaper most frequently read. Progress through our usual educational institutions is no index of increasing liberalism, according to the observations of Symonds (12, 1925); but Jones (28, 1926) found that, while the average score did not change much, the seniors, as compared with other classes, had shifted some of their conservatism out of the area of religion into the area of economics. One of the best studies of the correlation sort is Vetter's study (175, 1930) of social attitudes among students, which found that the radicals were more apt to be men than women, poor than prosperous, older children, nonpartisans, Jewish, and above the conservatives in intelligence. Allport's very extensive collection of student opinions at Syracuse likewise showed an advantage for the radicals in intelligence. The influence of culture on some attitudes is neatly demonstrated in studies by Anderson and Davis (33, 1927), showing how differently occupations are rated in social status in the United States

and in Russia. Nationality differences in some attitudes were further reported by Abel (28, 1926) and sex differences by Lundberg (28, 1926).

Correlates of religious attitudes were studied by Van Ormer (33, 1927), Starbuck and Sinclair (33, 1927), Howells (30, 1930), Ewing (36, 1929), Watson (18, 1929) in a study of worship preferences, and by Woodward in a still unpublished study of the relationship between adult religious patterns and the type of emotional relationship existing between the individual and his parents during childhood. The University of Iowa studies (Shuttleworth, Sinclair, Howells, all with Starbuck) show rather consistently that individuals from religious homes, with mystical experiences of God, conservative in religious beliefs, tend to be less intelligent, and more suggestible.

The most convincing method for the study of attitudes and opinions is not through correlations but through tests given before and after an experimental process designed to bring about a change. College classes in social science have been tested by Harper (169, 1927), with very encouraging evidence of change, but by Zeleny (28, 1926), Kornhauser (30, 1930), and Fowler (34, 1928) with less evidence for modification of attitude by the course and more, perhaps, for the stability of opinion measures over periods of time. Porter's excellent but unpublished thesis at the University of Chicago on pacifist and militarist attitudes was based on a test, the items of which were scaled in accord with the actual answers of persons rated by their friends at various points between one extreme and the other. Less militarism appeared in the Congregational and Methodist groups, as contrasted with some other denominations; the R. O. T. C. officer groups generally scored high in militarist attitude; but the effect on students of participation was not clear. Conferences and conventions have been studied by Watson (177, 1925; 36, 1929), Lamb (37, 1930), Wubben (36, 1929) and others. Witmer (18, 1929) found that a course in sex education did not greatly improve the attitude of mothers. Watson (111, 1928) found that two weeks in Y. M. C. A. summer camp brought about definite improvement in the average attitude toward law and discipline, but less significant, although still positive, changes in attitudes toward other nationalities, races, religions, and classes, and toward the meaning of camp for the life of a boy. Patterson (37, 1930) found that boys sent on good-will excursions to Japan commonly acquired more new information but little change in basic attitude. Sturges (16, 1927) found that reading an article produced a positive change, even when individuals were antagonistic to the article. Wheeler and Jordan (18, 1929) demonstrated the considerable influence on student opinion of knowing what most people think, or what experts think. Bird (33, 1927) found a given newspaper error capable of misleading even people who had experienced the original speech. Thurstone found that movies in which Chinese are heroes or villains do have a definite tendency to raise or lower,

correspondingly, the status of Chinese among other peoples mentioned in the list to be arranged in rank order. It is probable that in no other phase of character education do we have so many scientifically established facts; although general laws or principles can scarcely be derived from the present evidence.

#### Test Materials Now Available:

<i>Title</i>	<i>Publisher</i>
Allport, A Study of Values	Houghton Mifflin Co.
Bogardus, Social Distance Test	Bogardus. Also certain forms available from C.E.I.
Case, Test of Liberal Thought	Columbia University
Harper, Study of Opinions, Feelings, and Attitudes Concerning Some International Problems	Association Press
Harper, Social Attitudes Test	Columbia University
Hill, Test of Civic Attitudes	Public School Publishing Co.
Neumann, Test of International Attitudes	Columbia University
Opinions on International Questions	Association Press
Opinions on Race Relations	Association Press
Religious Thinking Test (Elementary or Advanced)	Association Press
University of Chicago Attitude Scales—God, church, war, Negro, birth control, with 25 or 30 others in process.	University of Chicago Press
Watson Test of Public Opinion on Religious and Economic Questions (Fair Mindedness)	Columbia University

#### Originality, Imagination, Resourcefulness

The close relation between these qualities and the tests of the early psychological laboratories and of intelligence, brought relatively early development in this field. In 1898 Dearborn (12, 1898) reported on a study of imaginations. Interference and adaptability were reported by Culler (12, 1912). In 1916 Chassell (178, 1916) outlined a number of tests related to originality, but McClatchy (34, 1928) found, as has so often happened with tests of supposed character qualities, that the intercorrelations were distressingly low. Following up the early work on imagination, using some tests suggested by Whipple (12, 1915), McGeoch (12, 1924; 180, 1924) again found rather low intercorrelations. Lundholm (12, 1924) studied imagination in relation to mental disease, Simpson (12, 1922) in creative activity, Teague (12, 1922) in music. Deutsch's test of conformity (12, 1923) takes rank with the most original tests in the field of character measurement. Subjects are asked to choose the best expression in proverbial form, the most comfortable living room, the prettiest girl, the best idea of the hereafter, etc. The alternatives are chosen from various civilizations and cultures. The score depends upon the



extent to which the individual finds his own culture always the best form. Struve (30, 1930) tried out, on two hundred cases, tests of imagination from ink-blot, from incomplete stories, and from stories built around stimulus words, and found considerable consistency and agreement with ratings. Resourcefulness in the chemical laboratory was studied by Beauchamp and Webb (33, 1927) with actual conduct tests. Given some elementary materials pupils were required to find ways of producing the desired results. The Bureau of Public Personnel Research also has developed tests of resourcefulness as reported by O'Rourke (18, 1929).

### **Perseveration**

The perseverative factor is named "p" in Spearman's outline (183, 1927). Another discussion of the psychological phenomenon of perseveration may be found in Lankes' review (182, 1915). Still unpublished data of Stephenson's at the University of London seem to show that the problem pupils in a class may be defined with unusual certainty (the claim is about 80 percent) in terms of those exceptionally low or exceptionally high in tests of perseveration in such simple psycho-motor functions as making x's, or crossing out e's. The perseveration appears in these tests when the task is slightly altered, e.g., to making plus signs, or crossing out a's. The preliminary reports are especially challenging, since no other simple tests promise anything like this degree of identification of character problems. In America very little attention has been given to this problem, the outstanding exception being Cushing's study (181, 1929) of pre-school children in natural situations. Intercorrelations of five tests averaged .42, which suggests more consistency than is usually found in a so-called character trait. Perseveration is defined by Cushing as a tendency to continue a task after external pressure has been reduced to a minimum. The fact that the trait is more psychological and less ethical in origin may contribute to the better promise of these early studies.

### **Persistence, Perseverance, Effort**

One of the remarkable pioneer tests was Fernald's (186, 1912) investigation of the length of time an individual would stand on his toes, without support. He found that the physiological limit was rarely approached, but that as the task grew increasingly distressing delinquents gave up quickly, while normal and successful individuals continued. Chapman (185, 1924) measured effort in a simple and monotonous task, studying its relation to speed in influencing success. Feingold suggested a measure of effort among high-school pupils. Watson (111, 1928) used the omissions on a more or less voluntary test performance as a measure of the amount of effort which the camp administration could call forth in boys. Morgan and Hull (187, 1926) used an alley maze, which could be made increas-

ingly and infinitely difficult. The Porteus maze has possibilities as a measure of stick-to-it-iveness, as have many other performance tests, notably the more difficult ones like the Ferguson form boards. Poull and Montgomery (18, 1929) report that the Porteus maze measures in delinquents something which the Stanford-Binet does not catch, and which discriminates the delinquents, to some extent, from normals. Lewin (30, 1930) and his students have used a number of monotonous tasks as instruments for the study of the total personality reaction of the individual. *Psychic satiation*, they find, is greatly influenced by factors inside the particular task and factors in the general environment and life of the individual. The stronger the affect attached to an activity the quicker repetition produced satiety. Students grew quickly bored at filling pages with short pencil strokes, but unemployed, who find a comfortable warm place to sit, with pay, continued indefinitely. Such illustrations call attention to a common fault of any persistence test which is defined purely in terms of a task, and which does not also prescribe the structure of the inner state and the other dynamic aspects of the present situation.

The most extensive tests in persistence are the Hartshorne-May series (71, 1929). One of their tests was constituted by a scoring of the Maller Cooperation Test, not by comparison of work done for self and for class, but by comparison of work done near the beginning of the period with work done on this boring addition exercise near the end of the period of work. Another test, the Story Resistance Test, was made of a story read aloud up to an exciting climax, with the ending badly pried in the printing so that only the more persistent children would carry through to find out the end of the story. Most puzzles that are difficult may be regarded as persistence tests. Some mechanical puzzles and some paper and pencil tasks were used in this fashion by the Character Education Inquiry. A series of individual tests given to twenty-five orphanage children by Hartshorne and May included: picking up pennies and ballbearings scattered over the floor (score was time until subject gave up, decided he had found enough of them); time subject would stand on right foot; time spent working for a dime, visible in a difficult puzzle setting; time from beginning to eat a cracker until subject can whistle or until he reaches for water to help him clear his mouth more quickly. Intercorrelations were close to zero, correlations with ranks about  $-.3$ . In the tests more widely used (stories, puzzles, and Maller exercises) intercorrelations averaged  $.24$ , and the correlation with reputation was  $.23$ . The very considerable influence of the administration and class-setting, the suggestion influence of one upon another, is evidenced by a correlation of  $.74$  between the persistence score of the average boy and the average girl of the classes taken as units. Correlation of siblings was  $.40$ ; correlation of persistence with age was  $.30$ . Low occupation levels did better on persistence tests than on tests of other character traits in this investigation, possibly because the puzzles

represented materials which the underprivileged youngsters attacked with more zest than they found for the other tasks. Correlation of persistence total with general all-around character was .44.

#### Test Materials Now Available:

<i>Title</i>	<i>Publisher</i>
Fernald Achievement Capacity Test	Stoelting
Maller Test—Series P	Association Press
Stories and Puzzles Test—Series P	Association Press

#### Physiological Indices of Character

In addition to the morphological and structural factors previously discussed, many other physiological symptoms are supposed to be reflected in personality and character changes. Hyper-thyroid means not only changed heart rate but also changed excitability; Graves' disease brings not only tremors but worry and anxiety. Failure of sex glands to develop means not only structural but also emotional infantilism. These diagnostic procedures are not quantitative in the usual test sense, but they are more valid and more valuable for guiding future procedure than most tests which deal with the behavior symptoms directly.

A more specific factor which has been studied is the acidity or alkalinity of the body fluids investigated by Rich (91, 1928), Starr (33, 1927), and Robertson (33, 1927). Rich found creatinine content similarly an index of excitability. Ferrari (44, 1928) found an increase in erythrocytes due to examination strain, as large as 457,000 per cubic millimeter. The recent interest in blood-group types led Furukawa (30, 1930) to investigate a possible relation to temperament, and Proescher and Arkush (188, 1927) to conclude on the basis of very extensive statistics that persons of type IV have only about one-fourth the liability to psychosis found in types I or III. Many investigations of physiological factors related to psychoses have been made and need not be listed here. Schizophrenics can be found to have almost any postulated sort of disorder, heart, digestive, gonadal, respiratory, reflex, etc., more commonly than a control group of normals. The significance of the inferiority is supposed to vary with the history of the individual case. Whitacre and Blunt (33, 1927) found that digestion is not so uniformly an index of disposition among normals as it is commonly supposed to be. History of disease plays the major role in the investigations of Stratton (189, 1926) on anger, Mühl (3, 1923) on tuberculosis, and Notkin (3, 1928) on the relation of early childhood diseases to epilepsy; each study yields some positive correlation between disease history and present makeup. Mention may be made here, also, of the Jaensch type studies (266, 1929; 265, 1930; 267, 1930) grounded on a supposed connection with thyroid and parathyroid activity. Paterson's

recent book (30, 1930) on physique and character summarizes most of the evidence and comes to the conclusion that few of the supposed connections are of real significance.

### Psychogalvanic Responses

The several hundred studies listed in the bibliographies on the psychogalvanic reflex are representative of more intensive work than has been given to any other similarly narrow problem in the field of character and personality measurement. Beginning, perhaps, about 1888 with an article by Fréré (12, 1888) the studies were so common as to justify a book in 1908 by Veraguth (193, 1908). They may be classified as studies concerned with the nature of the physiological process, studies concerned with reliability and technic, studies concerned with the reflex as an indicator of emotion, and finally miscellaneous applications of the reflex.

Early studies concerned with the nature of the process include those by Sidis and Kalmus (12, 1908), Müller (12, 1909), Radecki (12, 1911), Albrecht (12, 1910), Leva (12, 1913), Gildenmeister (12, 1913), and Aebly (12, 1919). Richter's analysis (192, 1927) of a case of unilateral sweating due to a lesion in the sympathetic system gave an excellent opportunity to show that the reflex was dependent upon the functioning of the sympathetic system. He concluded that the first short change might be produced by the sweat glands, that the slower phase went with deeper changes. Sweating produced by pilocarpine was accompanied, however, by increased resistance. James and Thouless (28, 1926) gave an explanation in terms of polarization.

Reliability was investigated by Sidis (12, 1910) and by Bartlett (33, 1927), who found that the type of curve typical of a state like anxiety or shock differed from individual to individual but remained constant over at least a three month interval for a given person and emotional state. Cattell (34, 1928) found resistance constant for a given individual and attitude, independent of ordinary fluctuations in temperature and humidity. Wechsler and Jones (34, 1928) reported self-correlations for the situation of effort (.81), aggression (.80), and startle (.67). For similar stimuli they found correlations of about .50 in the resulting deviation; for different stimuli the correlations were very low. Reaction to a word depended not only on the word itself but also on its position in the series. Compact sets of apparatus for recording the phenomenon have been made by Wechsler (12, 1925) and by Hathaway (17, 1928). Malmud (34, 1928) found that original resistance of the skin correlated .84 with the extent of the reflex reaction. Jones (30, 1930) tried out the test on babies, and found that babies age three to eleven months reacted in much the same form as did adults, although they had a lower initial resistance.

The main interest in the phenomenon as a character test has rested upon the fact that emotional changes often are paralleled by psychogalvanic reflex reactions. Many studies, including some by Peterson (12, 1907), Pieron (12, 1910), Waller (12, 1917-19, 1921), Binswanger (12, 1919), Smith (87, 1922) and Wechsler (12, 1925), pay particular attention to this phase of the problem. Syz (28, 1926) found that there was little agreement between subjective reports of emotional affect and the galvanometer, but concluded that the reports were modified in the direction of having "proper" emotions at proper times. One important fact is that conscious effort, as in problem solving, calls forth the typical psychogalvanic response. Gopalswami (28, 1926), Abel (30, 1930), and Rackley (30, 1930) found that the disturbance so caused is less than the response to fear. Malmud (34, 1928) found the largest response accompanying the irritation which came when attempts at ball-tossing were disturbed by a magnetic pull on an iron cuff. Bayley (190, 1928) showed that the curve for "startle" differs from that for "apprehension." Paterson's review (30, 1930) led him to believe that no characteristic psychological experience can be identified with any characteristic of the curve. Landis (30, 1930) summarized the whole matter very capably, and concluded that the psychogalvanic reflex is one of a series of related responses linked with the sympathetic nervous system, but that it may appear with no emotional accompaniment, or it may not appear when emotion is quite definitely present. He believes it worth studying as a phenomenon, but not as a test of emotion.

Psychogalvanic reflex technics have been used in the study of deception by several writers; of psychopathic states by many more; also in the study of drug influences by Waller (12, 1919); multiple personality by Prince and Peterson (12, 1908); indecision by Wechsler (12, 1922); emotionality during revived emotions by Washburn and others (28, 1926); preferences by Sastry (15, 1926); magnetic personality and nervous temperament by Fleming (15, 1927); strength of instincts by Collman and McRae (33, 1927); the relationship of sleep and hypnosis by Estabrook (30, 1930); and, by recording the responses to ambiguous words and phrases, as a test of "pure-mindedness" by Brown (12, 1925).

### **Ratings, Reputation Measures**

It is not hard to select evidence to make a case for or against this most common method of character appraisal. Those who are critically disposed toward ratings will point, of course, to Rugg's heroic study (205, 1922) of men in officers' training camps during the war, and to the fact that he could get all degrees of disagreement about men who were exceedingly well known to one another, that the agreement between objective intelligence tests and ratings on intelligence was not high, and that all in all he felt that a "single judgment by a single school officer can rarely be expected to place a child within his proper fifth of the total group." They will point,



further, to many other studies in which rating reliabilities have been low, for example, to Remmers and Plice (28, 1926), who found the carefully devised Purdue scale yielding reliabilities between .0 and .5. Arlett and Dowd (28, 1926), to cite one more example, studied ratings given by leaders to girls in summer camp. The girls were intimately known after two months of living together, but it was not impossible for the same girl to be ranked among the top 10 percent by one leader and among the bottom 10 percent by another leader on the same trait. They may further emphasize Newcomb's finding (134, 1929) that leaders "read in" to their observations of pupils relationships which are not to be found when the observations themselves are collected, or may, perhaps, go back to Thorndike's famous presentation (12, 1920) of the "halo error," whereby individuals tend to be rated on all traits in accord with a general approval or disapproval felt by the rate.

On the other hand, it is easy to find studies in which rating reliabilities run up to very satisfactory heights. Furfey (28, 1926) secured ratings on social maturity which showed reliabilities of .90; Hughes (28, 1926) found after a year that 53 percent were placed in the same fifth of the scale in which they had formerly been placed; Autenreith (17, 1928) found the agreement between two teachers who had been with twenty-eight pupils for three years as high as .89; and Bridges (34, 1928) found the consistency of ratings on pre-school children throughout a year averaging .78.

The outstanding factor in raising the prestige of ratings has been the Character Education Inquiry, in which Hartshorne, May, and Shuttlesworth (146, 1930) had opportunity to compare the effectiveness of what they happily named "reputation measures" with other measures related to character. The technics used in reputation measurement were as follows:

(1) *The conduct record* described several degrees of each type of behavior in multiple-choice form, e.g., reliability, cooperation, open mindedness, etc. The observer, usually the school teacher, checked the phrase which described the usual response of the child in each trait area. The correlation with test of cooperation was .12 to .28, with persistence tests .05 to .12.

(2) *The checklist* was a long list of adjectives denoting desirable and undesirable attitudes and behaviors. The teacher checked only the words which applied to the particular child rated on the blank. Correlation of list in positive form with list in antonym form was .74; correlation of two teachers .48. Correlations with service tests ranged from .03 to .44, with persistence tests from .00 to .33.

(3) *The Guess Who Test* was a series of brief pen-portraits of character types. The pupils wrote under the description the name of any classmate who seemed to them to correspond to that type. Reports were unsigned. Divided in two halves, reports gave scores with reliability of .88. Correlation between vote from pupils and vote from teachers was .80. Correlation with conduct test scores varied from .15 to .37.

(4) *Portrait matching device* was a measure of helpfulness or cooperation. Based on the pen-portrait idea but prepared for teachers to use in assigning to children numerical ratings, it is thus a little like the sample scales in use in judging handwriting or compositions. Correlation of scale used once and again by same teacher was .84.

Intercorrelations of these reputation measures with one another ranged from .22 to .63 with an average about .40. Agreement of all reputation measures of a trait combined, with all conduct tests of that trait combined, in the case of service was .50; of persistence, .23; of inhibition, .40. (These are uncorrected for attenuation.) Correlation of total reputation with school marks was .41, with school deportment .47, with emotional stability on the Woodworth-Mathews Test .38. A sex difference appeared consistently: girls were slightly superior to boys on all of the tests except the honesty series, but the girls were overwhelmingly superior on all reputation measures. Correlation of reputation in general with honesty tests was .18, with cooperation tests .28, with persistence tests .12, with inhibition tests .21, with moral knowledge tests .22.

The surprising contribution came when sixty-three judges, students of character education, were given one hundred portraits to judge. The portraits contained all the information, test scores, rating results, etc. about the pupils. These might be combined in any fashion which seemed to the judge best, to lead to a judgment on general all-around character. The resulting scale of portraits was used as a criterion against which to correlate each of the measures. In this situation reputation measures appeared to have been given much more weight than any other type of evidence. Correlation of the conduct record with general all-around character was .72. The checklist agreed with the criterion as indicated by a coefficient of .66; the coefficient for the Guess Who Test was .59. The general composite relationship of .61 between total reputation and all-around character may be compared with a relationship of about .45 for moral knowledge tests and slightly less for conduct tests.

Thus Hartshorne and May appear to have found that if reputation measures are taken both from teachers and children, by the better technics today available, they rival the best character tests in reliability and surpass them in relationship to the kind of character these judges were willing to approve. The relatively slight agreement between reputation and conduct tests must be interpreted in terms of the unreliability and possible lack of psychological validity in the conduct tests as well as in the reputation measures. In this connection it is interesting to note that Berne (195, 1930) with somewhat better relationship between test and rating situations found an agreement of .76 between the two types of measure applied to young children.

The question is thus shifted from the form, "Are rating scales good character measures?" to the form, "What kind of ratings, under what conditions, may be regarded as good measures of character?" There are several summaries of rating technic, notably by Norsworthy (202, 1908), Hollingworth (67, 1922), Knight and Franzen (199, 1922), Paterson (203, 1923), Hughes (197, 1925), Kingsbury (198, 1925), Watson (32, 1927; 29, 1927; 209, 1928) and the American Council on Education (204,

1928). These commonly stress such provisions as: (1) simple, clear, unambiguous description of behavior to be rated; (2) opportunity for many judges to observe the conduct in question; (3) increase in competence of judges through training, comparison and discussion of ratings; (4) selection of the traits and judges found to be more trustworthy; (5) improvement of the scale so that its units are more evident and more nearly equal; and (6) precautions against the "halo" effect (208, 1918).

Among the more specific and less generally emphasized contributions to technic are the following:

1. Poor performance tends to be more commonly noted and more reliably rated than good performance (202, 1908; 205, 1922; 198, 1925).

2. Close friendship tends to produce disagreement with the rating given by more casual acquaintances, and deviation is in a direction favorable to the person rated, as shown by Shen (12, 1925) and Knight (12, 1923).

3. No essential difference is found between the validity of the order of merit method and the value-assigning method so far as reliability is concerned, the latter being more congenial to raters, as shown by Stenquist (12, 1920), Conklin and Sutherland (12, 1923), and Symonds (12, 1925).

4. Seven intervals in a scale seem to be an optimum number, as shown by Symonds (12, 1924) and Curtis (12, 1923).

5. Judges may give reliable conclusions without being able to give good reasons for their judgments, as shown by Landis (12, 1925).

6. A general trait may be more reliably rated than a very specific one; but the combination of results from eighteen sub-traits is much more reliable than the rating on the one general one, as shown by Rugg (205, 1922), Slawson (12, 1922), and Furfey (28, 1926).

7. Judges take longer in rating disagreeable than agreeable traits, as shown by Dorcus (28, 1926).

8. Agreement between two judges is usually less than agreement between two ratings from the same judge at different times. The more widely separated the areas in which the two judges have opportunity to observe, the less agreement between them. All-around judgment demands judges from many different phases of the individual's life and social relations, but lack of agreement among them is in part a function of lack of consistency in individual living, as shown by Hanna (12, 1925), Watson (29, 1927), and others (200, 1927; 201, 1927; 209, 1928).

9. The best judges of others are rated by others as egotistic, cold-blooded, anti-social, as shown by Adams (33, 1927).

10. A relatively short scale of a few traits, say five or seven, will, because of the halo effect, serve about as well as one which is logically more complete, as shown by Kornhauser (200, 1927; 201, 1927), and Mort and Stuart (33, 1927).

11. Considerable time may be saved with little loss in reliable discrimination by asking judges to check only the more extreme or noticeable cases, leaving the great majority near the center of the distribution curve unclassified, as shown by Mort and Stuart (33, 1927).

12. Most judges tend toward leniency; they would like to rate all in the upper categories. Often the bottom part of a rating scale goes unused. This can be in part corrected by the use of terms which are not invidious and which provide distinct degrees of goodness. Further, correction may be made by the requirement that judges check a certain proportion of cases in each section of the scale. The most common correction

method, however, is to have the subjects ranked in order of merit instead of rated. Studies in this field are reported by Kneeland (34, 1928), Symonds (12, 1925), Hull and Montgomery (94, 1919), and Conklin and Sutherland (12, 1923).

13. Staggering the scale so that the favorable extreme is now at one end and now at the other should, theoretically, reduce the halo effect somewhat, as shown by Freyd (12, 1923), and Knight (199, 1922); but the results of one experiment do not confirm this expectation, that by Remmers and Brandenburg (33, 1927).

14. Signed ratings are less extreme than anonymous ones, as shown by Maller (30, 1930).

15. The degree of certainty which the rater feels, is positively correlated with the value of the rating and should be taken into account (78, 1923).

The scale blank itself is undoubtedly less important than many of the conditions surrounding its use. Nevertheless, some valuable suggestions have been made for improving such blanks. One of the early suggestions was Scott's Army Rating Scale, or Human Ladder, or Man-to-Man Scale (12, 1918), as it has variously been called. This scale has also been described by Terman (12, 1918), Rugg (205, 1922) and Paterson (12, 1922). Its characteristic feature was that outstanding individuals (the bravest, the most cowardly, an average man) were first identified with certain numerical values (e.g., 15, 3, and 9, respectively) and then other men rated by placing them, in imagination, alongside of these fixed points, giving a personal embodiment of what the figure was supposed to mean. One of the most common scale forms is the graphic rating scale in which check marks are placed along a line such as those by the Bureau of Personnel Research of the Carnegie Institute of Technology (12, 1921), Hayes and Paterson (12, 1921), Freyd (12, 1923), and Ream (12, 1921; 209, 1928). This yields quantitative results without facing the judges with the manipulation of figures. Hepner (28, 1926) found the adjective checklist better than the graphic scale, and Hartshorne and May (146, 1930), as noted above, followed this line. The Upton-Chassell Citizenship Scale (12, 1919, 1922) was an improvement upon most existing scales for school use because of its more objective description of schoolroom behaviors and its improved units. The best development along this line has been Hering's Scale (12, 1924) for measuring educational outcomes in terms of the ideals of the project method. He secured quantitative results, not by estimating the degree of interest, but by counting the number of pupils (or proportion of pupils) interested at the time of measurement. This translation into number of persons doing or not doing a fairly well-defined sort of behavior is a little like the methods which Thomas (60, 1929), Good-enough (136, 1928), Olson (58, 1929), and others are finding useful in behavior observation with short samples. Hering has recently developed for Y. M. C. A. experimental schools an improved form of this early scale. Yepsen's Score Card (34, 1928) is the preliminary form of what Hartshorne and May used as a conduct record. Symonds used the "Guess Who" technic in a new scale for measuring maladjustment in high-school pupils.

In addition to self-report measures (34, 1928) he has introduced a series of questions by means of which pupils may identify the classmate who is always fidgeting about, who is very easily upset over little things, etc. Using this dual report, from inside and outside the personality, he believes we may secure a better index of maladjustment than either alone would give. L. K. Hall suggested a rating technic in a still unpublished study which has further possibilities. Adjectives, phrases, or longer descriptions were placed on cards, one to a card. The judge then sorted the pack of cards, placing in one stack those which seemed to fit the person in question, in the other stack those which for one reason or another did not apply very well. Using the customary scale technic of Thorndike, Thurstone, and others, it is possible to have the phrases on the cards placed along a linear scale of whatever trait (e.g., general all-around character, or usefulness as a Y. M. C. A. secretary, or emotional maladjustment, etc.) so that each phrase has its proper quantitative weight as an index of the general trait. The scorer then averages the scale value of the phrase-cards said to apply to any individual to find that individual's place on the scale.

A notable contribution to thinking on scale and test construction is Adams' index (30, 1930) of subjectivity and objectivity in a measure expressed in terms of the ratio of self-consistency to group consistency. A highly subjective measure is one in which self-consistency is higher than group consistency among a group of judges using it. An objective scale is one in which agreement within the group is as good as agreement in the individual.

The uses of ratings have been manifold. So many studies have used ratings as a means of validating tests that it is impossible to recognize them all here. One comment will suffice. The practice of careful test construction, followed by hasty and inadequate rating technics, resulting in a low correlation between test and the rating criterion, leading to the dismissal of the whole piece of evidence because the ratings were not very good anyhow, is so absurd as to need no criticism. Mere description should reduce the number of such manoeuvres.

Ratings in the direct analysis of personality have been used by Heymans and Wiersma (262, 1906), Webb (184, 1915), Folsom (12, 1917), Garrett (28, 1926), Stead (28, 1926), and Hartshorne (18, 1929). Some of the results are discussed in the section on types and organization in character.

Ratings have been used as an aid in the selection, appraisal, and counsel of students by Moore (12, 1912), Chu (12, 1922), Rugg (12, 1921), Hughes (12, 1923-24), Alger (15, 1926), Earle (15, 1926), Kornhauser (200, 1927; 201, 1927), and the American Council on Education (204, 1928). Factors related particularly to school success were studied by Pressey (12, 1921), Poffenberger and Carpenter (12, 1924), Sangren (12, 1923), von Bracken (12, 1925), and Turney (33, 1927). Behavior



in the classroom has been rated by Plant (12, 1922), Chassell (12, 1924), Haggerty (12, 1925), Blatz and Bott (33, 1927), Wickman (210, 1928), and Stabler (19, 1929). The Haggerty-Olson-Wickman Scale (30, 1930) and the Merrill-Palmer Scale (30, 1930) have been used extensively.

Ratings have been applied to the study of gifted children by Burks (12, 1925), by Terman and others (138, 1925-30), and by Lamson (30, 1930). Ratings applied to those of inferior ability may be found in many studies, including those of Anderson and Leonard (12, 1918), Potter (12, 1922), and Porteus (12, 1920).

Ratings have been the most common method of measuring teaching personality. Among the better known studies are those of Ruediger and Strayer (12, 1910), Boyce (12, 1912, 1915), Connor (12, 1920), Rugg (12, 1920), Edmonson (12, 1921), Thompson (12, 1921), Wagner (12, 1921), O. M. Jones (12, 1921), E. S. Jones (12, 1923), Mead (12, 1922), Knight (12, 1922, 1924), Whitney (12, 1924), Crabbs (12, 1925), Schutte (15, 1926), Hamrin (33, 1927), Clem (30, 1930), and Light (30, 1930). Most of the studies are agreed that there is a something, a teaching personality, which is not indicated by intelligence or school marks, but which is vital for teaching success. Poise, address, sympathy, understanding, and cooperativeness, are among the terms often used. Crabbs (12, 1925) found little relationship between such judgments and the accomplishments of children on standard subjectmatter achievement tests, a fact which may be interpreted to the discredit of either of the measures applied. Hamrin (33, 1927) found little relationship between the ratings given students in practice school and ratings given them by supervisors in the field later, but the usual correlation is moderate. Schutte (15, 1926) and several of the others suggested that the scale may be used as an instrument of supervision to help the teacher develop his weak points. Ratings by pupils asked to judge their teachers were studied by Bird (12, 1917), Dolch (12, 1920), Guthrie (16, 1927), Stalnaker and Remmers (207, 1928) and Boardman (30, 1930). Boardman found that pupils, fellow-teachers, and supervisors each might contribute a special aspect of judgment, but that they agreed in their appraisals as indicated by an average inter-correlation of .6. Instructors rated at Purdue, in the Stalnaker study, when judged by ninety-four students were placed with a reliability of .73 to .96, depending a little on the trait. This means, clearly, that student judgment of teachers, because of the larger group, is more reliable than teacher judgment of students, and, indeed, more reliable than most standard tests. It would seem to be a factor well worth taking into account.

Almost every personnel department in industry has at one time or another used rating measures in relation to success. When the person who does the rating is responsible for the promotion policy, there is apt to be a very close correlation between ratings and promotions. This is like the correlation between ratings of the school teacher on "application" and the

teacher's marks. Both contain the common factor of subjective bias. Achilles (12, 1917) and Kohs and Irle (12, 1920) made early studies of the relation between trait ratings and the fact of promotion in the army. The Bureau of Personnel Research of Carnegie Institute of Technology (12) developed in the years immediately following the war a variety of scales related to clerical, sales, executive, and other types of business success. Gallup (28, 1926) found that a rating scale was preferable to mental tests, tests of memory of trade-marks, introvert-extrovert tests, interest tests, or tests of supposed social intelligence, in predicting success among retail salespeople. Such studies are typical of hundreds of others, published and unpublished.

Among the miscellaneous uses for ratings may be mentioned the study of race differences by Davenport (12, 1921, 1923), Murdoch (12, 1924), Porteus (12, 1924), and Porteus and Babcock (28, 1926); the investigation of characteristics of persons achieving distinction in science by Cattell (12, 1903, 1915); the study of nervousness and sleep by Terman and Hocking (12, 1913); the differentiation of socially and mechanically minded persons by Freyd (12, 1922, 1924); the study of the relation of social maturity to physical and mental maturity by Gates (140, 1924); sex differences by Hart and Olander (12, 1924); emotionality by Landis, Gullette, and Jacobson (86, 1925); accident liability by Payne (12, 1923); social adequacy by Porteus (12, 1920); correlation between character and I. Q. by Chassell (28, 1926); desirability of traits by Yoakum and Manson (28, 1926); traits of homemakers by Charters (15, 1926); relation of personality to morphology by Sheldon (33, 1927); effect of family relations on personality by Goodenough and Leahy (33, 1927); success in club and camp leadership by Statten (18, 1929), Dimock and Hendry (196, 1929), Ure (37, 1930), and Bartlett (36, 1929); and qualities of student leaders (135, 1926).

#### Test Materials Now Available

<i>Title</i>	<i>Publisher</i>
Behavior Frequency Scale, Form B	Association Press
Camp Behavior Frequency Scale	Association Press
Checklist (C.E.I.)	Association Press
Colgate Scale for Measuring Executive Leadership	Hamilton Republican
Conduct Record (C.E.I.)	Association Press
Fundamentals of Character Rating Scale, Form F	Association Press
Guess Who Test (C.E.I.)	Association Press
Haggerty-Olson-Wickman Rating Scale	World Book Co.
New York Rating Scale for School Habits	World Book Co.
Portrait Matching Device (C.E.I.)	Association Press
Situation Rating Scale, Form S.	Association Press
Trait Rating Scale, Form T.	Association Press
Upton-Chassell Citizenship Scale	Columbia University

## School Success and Failure: Character Factors

The well-known correlation of about .5 between grades in high school or college and the results of intelligence tests leaves three-quarters of the factors influential in producing academic success unmeasured by intelligence tests. It is natural, therefore, that school leaders should experiment with existing character tests and should devise new ones particularly designed to account for the remaining factors. Toops (218, 1927) suggested some of the phases of character and environment which may be worth investigating. An early study in this direction was May's (214, 1923) pointing to a negative correlation between study-time and grades. Kauf (33, 1927), McCabe (33, 1927), Newcomer (33, 1927), and Sturtevant and Strong (33, 1927) are among the others who have used time-schedules to measure a factor related to school success. The able pupils in general studied less and spent more time in extra-curricular activities than did the students who received lower grades. Stoke and Lehman (30, 1930) found poor students more apt to exaggerate their report of study-time. A less objective and less revealing method of study has been the collection of character ratings on successful and unsuccessful pupils. Pressey (12, 1921), Sangren (12, 1923), Hughes (15, 1926), Flemming (15, 1926), Turney (33, 1927), Adams, Furniss and DeBow (34, 1928), Steere (18, 1928), and Herriott (212, 1929) have followed this line. The results show, of course, that the pupils to whom teachers give good marks are also given by the same teachers, ratings which those teachers believe would justify their good marks. "Good" students are, almost by definition, those who are regarded by their teachers as accurate, sensible, conscientious, resolute, mentally "quick" and "deep," ambitious, industrious, perseverant and the like. A third approach has been through the application of such tests as the Downey by Poffenberger and Carpenter (12, 1924), Kolstead (12, 1924), Miner (12, 1925), Reaves (12, 1925), Stone (12, 1922), Downey (12), Oates (34, 1928), and others; or the Pressey X-O by Pressey (12, 1921) and Chambers (12, 1925). As indicated previously in discussing those tests, the results have not been promising.

Symonds (12, 1925) first proposed that studiousness be measured by the relationship between intelligence and test performance on assigned material; next (28, 1926) proceeded to observation of the actual behavior of studiously successful and unsuccessful pupils; and then experimented with interest questionnaires (217, 1928) to discover whether academic prowess might be predicted on such a basis. He found that most of the study-habit doctrines did not appear significantly to differentiate these groups, but that the better students were more apt to raise questions, to hand work in on time, and to work on to the very close of the study period. The interests characterizing the boys who won good marks in high school were of a rather negative and "sissy" type: less interest in revolvers, aviation, varsity teams, visit to moonshine stills, becoming a short story writer,

owning poker chips, owning a bull dog, etc., more interest in study, keeping rules, passing mark for eligibility, condemning betting, attending lectures or art museums. Kornhauser's (33, 1927) experience was typical of what usually happens with interest tests in this situation. From a large list given to 110 freshmen he selected 114 items differentiating successful from unsuccessful students. The correlation between test score and academic success in the group upon which the items were selected and weighted was .73. Administering it to a new group, however, the correlation between test score and academic success dropped to .17. Obviously one administration or even two, may be assumed to show differences on a chance basis, which makes the selection useless for further groups. Shuttleworth's study (216, 1927) is the best of the interest analyses. He found, as did Kornhauser, that his like-dislike test fell from a correlation of .63 the first year to one of .09 when applied the next year. Certain items of interest, however, remained constant. In general the good student at Iowa showed less liking for mechanics, were more unconventional, more independent, more critical of orthodox religion and patriotism, more tolerant of other groups and races, less worried, less confused, participated more in forensics, and had more cultural interests.

Studies by Remmers (215, 1928) and Fleming (34, 1928) are characterized by the use of a considerable battery of measures. Fleming found that none of the emotionality and neurotic-symptom questionnaires had any appreciable relation to the success of students at Columbia. Remmers confirmed this, but found differences five or more times their P.E., indicating that the successful students were well rated by their high-school teachers, were judged subjectively by the investigator to have a proper motivation, and to come from the city rather than the country.

The most thorough study of this problem was that by Gladys Watson (219) in which experienced educators were rated by several faculty persons who knew them well, as likely or unlikely to achieve marked professional success. A large battery, including the Strong Vocational Interest Test, the Stanford (Jensen) Educational Aptitudes Test, the Chassell Experience Variables, the Thurstone Personality Inventory, the Harper Social Attitudes Test, and two tests of her own, one of characteristic responses in educational situations and another of information about modern social, political, scientific, and cultural life, was administered to each person. In addition, intelligence tests, school marks, and time-schedules were available. In some cases interviews were added. Results were analyzed not only in terms of test scores but also for each of the thousand or more item responses. Her conclusion was that no test or type of test was likely to prove of value for this discrimination; that only the qualitative analysis of the relationships in each particular case gave any insight into the probability of success or failure.

Other scattered studies suggest that smokers are less apt to succeed, e.g., Earp (33, 1927); that fraternity and especially sorority members excel "barbs," e.g., Eurich (33, 1927); that better students are younger, lighter in weight, shorter, and healthier, and do less outside work, e. g., Kauf (33, 1927) and McCabe (33, 1927); that intellectual introverts provide few failures while extroverts provide more, e.g., Young (33, 1927); that introversion has no relation to success, e.g., Guthrie (16, 1927); that the Woodworth, the Pressey X-O, and the Kent-Rosanoff tests are not much help, e.g., Bridges (16, 1927); that the Woodworth test and social intelligence tests are not much help, e. g., Peatman (34, 1928); and that there is some relationship to economic status, e. g., Chauncey (211, 1929).

The motivation leading to effectiveness in school has been approached through what may be called test-experiments, and these seem more promising. The most common experiment has been to administer praise to one section, and reproof or disregard to others, as evidenced in studies by Gates and Rissland (12, 1923), Gilchrist (12, 1916), and Hurlock (213, 1924). As a rule these show the superiority of approval in general, but have not been used to show individual differences. Ross (33, 1927), Sullivan (33, 1927), and others confirmed Thorndike's finding and thesis that activity followed by success is more likely to be repeated than activity the results of which are unknown or which is found unsuccessful. Lewin (275, 1926-32) formulated the law of reaction to success and failure in other terms, in accord with a different psychological viewpoint, and suggested that the type-reaction to failure (try harder, give up, grow angry, etc.) is constant for the individual through many situations, although his data are not in a form to permit comparison with other attempts at character testing. Other motives of significance for school work are involved in Knight's study (12, 1922) of unwillingness to be tested; Sims' study (34, 1928) of competition; Kendrew's study (30, 1930) of the strength of desire for food, curiosity, and competition in young children; and Leuba's study (30, 1930) of rivalry, praise, recognition, desire for sweets, etc., in helping learners move beyond the plateau. A slight modification might make many of these studies significant for differential psychology as tests.

### Self-Appraisal

Self-rating is sometimes, very naïvely, supposed to be a means for finding out the individual's real persistence, cooperation, emotionality, etc. When more wisely used these ratings are regarded as measures of what the individual reports about himself, how he conceives himself, what his self-insight may be. A suggestion by Knight and Franzen (199, 1922; 220, 1924) was developed by Watson (29, 1925), Tyler (50, 1930), and Sweet (30, 1930), as discussed above in connection with disguised measures for detecting emotional abnormalities. In principle it involved a series of compari-



sons between an individual's idea of himself as he is, his idea of himself as he ought to be, and his idea of the average person with whom he might be compared. Shaw's study (223, 1931) is the most thorough exploration of the self-appraisal problem. He compared students' expectations of success in class with their grades, their rating of achievement with standard test scores, their ideas of their strong and weak points in practice teaching with the ideas of their supervisors, and their insight into themselves as indicated by recognition of common adjustment patterns. His findings included the usual one, duplicated with almost every "trait" test, that self-insight in one direction was no basis for prediction of self-insight in some other matter.

More limited studies by Shen (224, 1925) and Yoakum and Manson (15, 1926) show that the reliability of self-ratings is fully as good as, perhaps slightly superior to, ratings by others; by Hurlock (16, 1927) and Kinder (12, 1925) that there is a general tendency in our civilization toward over-rating of oneself on desirable qualities, especially marked among those at the low end of the scale; by Trow and Pu (16, 1927) that Chinese are less apt to overrate themselves; by Dorcus (15, 1926) that it takes longer to rate oneself than to rate a classmate. Others studied sex differences, for instance, Uhrbrock (225, 1926) and Heidbreder (30, 1930), and the correlation between self-rating and ratings given by others, with results usually in the neighborhood of .5, for instance, Washburn and Stepanova (12, 1923), Shen (224, 1925), Jackson (18, 1929), Flory (30, 1930), and many others. Uhrbrock (225, 1926) found murderers in a penitentiary rating themselves much as men students in college rated themselves and assumed that this showed some inadequacy in the ratings. Schutte's studies (222, 1928) show that on rating class work low and high students both tend toward moderation in rating themselves, the low students evidencing more distortion than the high. Maller found that children on unsigned ballots were much more apt to vote honors for themselves than on signed ballots (30, 1930). Meili (34, 1928) showed that people are very suggestible regarding their own characteristics: more than 50 percent of students accepted more than 50 percent of the trait characterizations supposedly attributed to them by a character analysis, but actually assigned by sheer chance. Laws (15, 1926) used self-ratings of parents as a means of parent-education. The technic has often been applied in teaching and supervision, for direct educational effect. The Find-Yourself Blank is an attempt to use such a technic in vocational guidance for high-school boys.

### Sex Differences

No phase of human life is more interesting than the polarity of sex in which the "unity is divided; the diverse unified." In addition to physical differences, determined in part by heredity and in part by the many factors

which may influence the development and interaction of the glands, there are apparently deep-seated psychic differences. The pure masculine is regarded as active, creative, productive, moving, willing, inspiring, powerful, dominant, objective; the pure feminine as receptive, birth-giving, passive, quiet, motherly, soulful, tactful, emotional, patient, sympathetic, altruistic, vivacious, artful, submissive, subjective. Experimental investigations of differences in ability were summarized by Lipmann (229, 1924) who concluded that men excel in weight discrimination, optical space discrimination, time estimation, tendency to exaggerate the time interval, precision and coordination of movement, richness of detail in drawings, mathematical aptitude and achievement, technical interest, drawing ability, historical ability, political inclination, practicality, business sense, ambition, striving for power, sexuality, laziness, courage, earnestness, untruthfulness, humor, reasoning, distractability, and intelligence. Women, he concluded, after eliminating conflicting and ambiguous studies, are superior in spacial discrimination along the skin, taste sense, hearing ability, color discrimination, tendency to minimize time interval, speed of decision, penmanship, manual dexterity, speed in arithmetic, linguistic talent, school marks, intellectual inclination, philanthropic inclination, religious interest, vanity, industry, good manners, cheerfulness, orderliness, demureness, truthfulness, impulsiveness, and constancy of attention. Allen's reviews (226, 1927; 227, 1930) are much more cautious, because they are based on better quantitative investigations. It is generally agreed, however, that what differences may remain, after extensive testing, can be attributed in large measure to cultural influences. Anthropological reports from other cultures leave little doubt of this. Nevertheless, tests showing the position of an individual as between "masculinity" and "femininity" in our present society have a genuine psychological interest.

Miles and Terman (230, 1929) used word association methods and interest indicators and determined inductively the answers given more often by men or by women. On this basis Terman (138, 1925-30) found that the gifted children were distinctly nearer a mid-point, the boys not so extremely masculine and the girls less extremely feminine than his controls. Other studies point toward the possibility of testing masculinity-femininity, although not developed into test form. According to Hartshorne and May (71, 1929; 146, 1930) an excellent reputation, not fully justified by conduct responses, would be an indicator of femininity. Tendencies to be hurt, hesitant, to worry, characterized the self-reports of women students, while men were more conservative and more outspoken, in Heidbreder's report (228, 1927). Willingness to admit bad traits would count one for masculinity, according to Hurlock (33, 1927). Conversation touching on money, business, sports, would score for masculinity; on women, clothes, or men, would score for femininity (122, 1927). Having played truant would score for masculinity according to Williams (33, 1927). Lehman

and Witty (33, 1927) in their many play behavior studies point to sex differences in attitudes toward school work, in preferred games, in fortune telling, in aesthetic activities, in reading books for fun; these differences can be arranged by ages. The Lehman questionnaire (see list of interest tests) can be scored for masculinity and femininity of interest. Handwriting differences can be recognized but it is difficult to define them objectively. Newhall (28, 1926) found 58 percent, and Kinder (28, 1926) 68 percent of several thousand judgments correct as to the sex of the writer. Experts claim much better proportions, but there is always a middle group, difficult to classify. Noh and Guilford (30, 1930) used the free production of one hundred words as rapidly as possible, and found that a larger proportion of verbs, abstract terms, names of implements, and occupational terms, characterized men. Weinland (30, 1930) found that sex differences in choice of proverbs were noticeable. Many studies, of course, have gone to contradict the current ideas of sex differences. Reference may be made to Valentine's evidence (17, 1929; 231, 1929) that intuitive judgment of character is not more trustworthy in women than in men.

### Sociability, Social Acceptability

Popularity can be very objectively measured by appeal to the group. Those children whom others wish to have as chums and companions are popular. The approximately one in ten whom Hartshorne found to be unwanted as a "best friend" by any of his classmates, is by that fact placed at a low point in the sociability scale. On the basis of chum-choices Almack (12, 1922) found that school children tended to choose chums similar to themselves in intelligence and chronological age ( $r = .5$ ). Wellman (28, 1926), behavioristically, counted the times children were seen together; he found among 113 pairs that similarity in scholarship was the rule, and that among boys similarity in height, I. Q., and C. A. were important. Furfey (233, 1927) showed in addition to the factors just mentioned, the importance of living in the same neighborhood (48 percent of the cases) and especially of being in the same school class (89 percent of the cases). A review of some reasons for friend-choice is presented by Rasey and her associates in Detroit (236, 1929).

The beginnings of social response can be observed as early as the twentieth day of life, according to Zoepfel (239, 1929). Bühler's studies (26, 1930) included facial recognition and social smiling as indices of social maturity. Loomis' study (57, 1931) in the nursery school was based on a count of physical contacts which could be classified as friendly or unfriendly, aggressive or receptive. Such measures would seem to show, not only gross differences in sociability, but something about the nature and direction of social expression. Berne (195, 1930) used interest in the group as one of the measures studied in a group of nursery-school children, and reported a correlation of more than .4 with M. A. Verry (238, 1925) gave another

study of social attitudes observed in the free play of the nursery school. Goodenough (136, 1928) applied a short-sample technic to several phases of social participation in the nursery school.

Among older children sociability may be rated, according to Lehman and Witty, by the proportion of individual or solitary activities in relation to the number of play activities shared by others. Reading, of course, makes up the bulk of the solitary time, and hence this index tends to be greatly influenced by intelligence and other factors affecting interest in reading.

Social intelligence (235, 1926) is the name given to what is measured by a test of ability to remember the connection between names and photographs, to give correct answers as to the best procedure in social situations, etc. The test correlates better with general intelligence tests, as a rule, than with measures of social participation. Gilliland and Burke (234, 1926) found that a questionnaire on social participation was a better measure than the recognition of photographs. Binneweis (34, 1928) used participation in extra-curricular activities as an "index of communal spirit." Chapin (232, 1926) offered an interesting technic for weighting the participation of an individual in organizations of which he is a member, attendant, officer, etc. Hewlett and Lester (34, 1928) used ratings by the dean of women on sociability, expressiveness during interview, etc., to compare with measures of introversion and with self-ratings.

Social insight is presumably one of the factors measured by the Sweet test. (See section on abnormalities.) Another approach to social adjustment, giving more weight to emotional factors, is Baumgarten's Test of Sympathetic Intuition (12, 1922). A review of the literature on social relations of children was made in 1927 by Shuttleworth (237, 1927).

#### Test Materials Now Available

<i>Title</i>	<i>Publisher</i>
George Washington University Social Intelligence Test	Center for Psychological Service
Lehman Play Quiz	Association Press
Guess Who Test	Association Press

#### Speed

The quick and the slow have been prominent in type-categories of characterology for two thousand years. More recently, since the beginning of laboratory work, speed of reaction, speed of association, speed of discrimination, speed of problem solving, and so on have been measured in studies too numerous to mention. The fundamental question here is the consistency of this index. Are the people who are quick in one situation likely to maintain that tempo in other activities? It is the assumption of many psychotechnical tests like tapping or reaction time, and of such tests as the

Downey Will-Temperament Test, that this tempo is fairly constant and that small samples will give an idea of the general trait. It is the observation of many a school teacher that a pupil may dawdle to an almost unbearable extent in algebra study, but move like lightning in a basketball game. The scientific evidence tends, although not unqualifiedly, to support this common observation. Bridges (243, 1914) found some consistency in speed of decision. Bernstein (12, 1924; 241, 1924) reported a general speed factor, analyzed by Spearman's methods, in tests of intelligence. Thorndike differentiated speed, along with area and power, as factors in intellect. MacFarland (3, 1930) found a fair consistency in speed, and a fair correlation with general intelligence. Braun (242, 1927) found a personal tempo in tapping, walking, lifting, etc., fairly constant ( $r = .44$ ) if not disturbed by suggestion or pressure of some kind. The evidence presented by Uhrbrock (97, 1928) is the most encouraging, a correlation of .8 between split halves of fifteen different speed tests. Tapping alone had a correlation of .68 with the criterion. Kennedy (3, 1930; 244, 1930), using both simple and complex mental reactions, found that speed, or as he called it "irritability," had an average intercorrelation of .45, but was unrelated to intelligence. Hübel (3, 1930) found consistency in speed of movement, association and reaction time, but no relation of these measures to speed in shift of attention. On the other hand, the results of Trow (246, 1925), Dowd (3, 1926), and Baxter (240, 1927) give almost zero intercorrelations among speed tests. Mace (33, 1927) showed that "natural" rate of work can readily be changed by practice. Chapman (12, 1924) and Klineberg (245, 1928) showed that success may be influenced by persistence and accuracy, which may be quite independent of, or even negatively related to, speed. Klineberg thinks that our civilization may have placed a special premium on speed, which makes our tests unfair, for example, to Indians, who, even if they have our language, have a different culture pattern.

Useful tests for individual testing are described in Bronner and Healy's *Manual of Individual Tests and Testing* (3, 1927). Some of the other character tests, e. g., the Maller Tests (see cooperation) or the speed tests (see honesty) can be used to indicate speed of work in limited activities.

### Suggestibility

In Murphy's excellent review of this topic, attention is called to the fact that the term *suggestibility* is used with many different meanings (250, 1931). Some tests of suggestibility are simply sensory illusions. Others involve a learned set which perseveres despite a little change in the situation. In some a prestige factor is present; in others it is carefully eliminated. Others involve willingness to change judgments under group pressure. In this, as in many of the other areas discussed, progress will depend upon more careful analysis before plunging into test-making.



Brown's (247, 1916) was one of the earliest samples of conduct tests. Some of his material was used and standardized by Otis (251, 1924) in a test which is now available, but which suffers from the lack of clarity of psychological analysis suggested in the preceding paragraph. Scott (12, 1910) and Town (12, 1916) represent early and limited samplings of the behaviors called suggestibility. Allport's study (12, 1919-20) of social facilitation, Whittemore's study (253, 1925) of competitive consciousness, Moore's study (12, 1921) of susceptibility to majority and expert opinion, and Wheeler and Jordan's study (34, 1928) of the influence of the group on the individual's ideas, all show individual differences in the extent to which an individual is influenced by his fellows. This may or may not be properly termed "suggestibility."

Marrow (34, 1928) used the influence of suggested color and thermal changes and the "Aussage" Test in which memory for the details of a scene is influenced by "suggestive" questions. In the "Aussage" Test the relation between intelligence and non-suggestibility is usually strong, in this case about .80. McGeoch (249, 1925) found susceptibility to the size-weight illusion not closely correlated with intelligence quotients (-.15), while the progressive-weights test, in which the essential factor seems to be perseverative response, showed a correlation of -.47 with intelligence. The two so-called "suggestibility" indices were as might have been expected, not related (-.01).

Hull (30, 1930) suggested a series of experiments in which susceptibility to hypnotic suggestion may be quantitatively measured. One, for example, involves a thread which indicates how far the individual leans when it is suggested to him, while his eyes are closed, that he is falling forward.

Negativism may be regarded as a kind of reverse suggestibility, but such studies have been reviewed under the heading of cooperation. Perhaps cooperation, too, is a form of suggestibility. Clearly any supposed measure of "suggestibility" must be interpreted in terms of the particular test used. The correlation between such tests and life behavior in other areas is not known. The nearest approach to such evidence—the correlation between suggestibility and delinquency shown by McGeoch (12, 1925), suggestibility and dishonesty shown by Hartshorne and May (108, 1928), and suggestibility and religious mysticism shown by Sinclair (252, 1930) and Howells (248, 1928), both at the University of Iowa—is all capable of interpretation in terms of the fact that suggestibility goes with inferior intelligence. Young's study (34, 1928) of suggestibility in race differences and Crane's similar study (116, 1923) with the "guillotine" suffered similarly from lack of control of possible differences in intelligence.

#### Test Materials Now Available

<i>Title</i>	<i>Publisher</i>
Otis Suggestibility Tests	Stoelting

## Types: Underlying Organization of Character

Many times in the course of this review it has been necessary to recognize the lack of consistency in trait measures. Yet anyone who has studied people knows that there is a uniqueness in each person which is consistent. One does not fall in love with a person conceived as a heterogeneous assortment of situation-responses. The person is known and experienced as a unit. Inconsistencies appear, of course, but are felt to be part of an inner whole and, indeed, with better knowledge of the individual, fewer behaviors are felt to be "out of character." It would be easy to multiply illustrations all pointing in the same direction to the conclusion that people are regarded as generally true to themselves, to the type of person they are, and not merely to the situation which stimulates the particular response. After all due allowance has been made for a tendency to unwarranted generalization from a few instances, to misjudgment of persons in accord with our preconceptions, there still remains a substantial body of observations supporting the idea of a unified personality. Fundamental to all character testing is the question of what this unity is and how it may be conceived. The tendency in American psychology will be to seek an inductive and empirical answer. Test a thousand behaviors and study their statistical interrelationship. So the answer runs. But it is no real answer. It misses the main point. What behaviors shall we test? How shall their likenesses be sought? Back of the collection of evidence must lie a sound theory of character derived from study of the whole rather than of random elements.

Absence of such analysis seems to characterize most of the present empirical data. Heymans and Wiersma (262, 1906) collected ratings on several thousand persons on ninety characteristics, arranged in accord with the theory that the fundamental differences were in emotionality, activity, and perseverance. From this they derived eight types, but the inclusion in the same type of such different persons as Michaelangelo, Pasteur, and Nietzsche shows the extraordinary inadequacy of these categories for giving us persons who seem to be "really" alike. Webb's all-embracing study (184, 1915) was made, not to find the organization of character, but to find some common element, which appeared to be volition, "w". McDonough (18, 1929) found that common factors of will, cheerfulness, sociability, and emotionality could be identified; but the very use of such terms and of the subordinate categories indicated an assumption about the unity of traits which her experiments could not check. Newcomb's evidence (134, 1929) points against such consistency. Hull's study (33, 1927) of variability in ability within the individual is useful in showing that the range of performances in a person may follow a normal curve if a chance series of measures is applied, but gives us no help in understanding how these are organized. Thorndike in his first volumes, Jersild's

study (30, 1930) recently, and many in between have shown a positive correlation among desirable qualities, but this is far too low to permit any assumption of uniformity on some scale of social values. The most naive proposal possible is that the individual organization be measured in accord with the expectation that he should compare with others as well on one trait as on any other. If his memory be a 60 percentile memory, then this very naive expectation would be that his will-power, his attention, his reaction time, his cooperation, and his religious aspirations should be 60 percentile, too, in order to make him a well-organized person. Mechanical as this sounds, it was actually promulgated by Garrett (28, 1926) and forms the basis for the study in organization of character made by Harts-horne and May (146, 1930).

Of greater promise is the attempt to study the consistency imposed by the environment. L. K. Hall in an unpublished study secured ratings on boys from many sources, not to average them in an "all-around" measure, but to analyze the difference between the boys who were uniformly rated in their various group participations and the boys who, in contrast, were given ratings which varied greatly from one situation to another. He found that the "problem" boy was usually characterized by wide disagreement among the judges rating him. Approaching the problem of variation in environment from its effective "inner" aspect, Spencer is working on a test of the subject's awareness that different standards are set for him by mother, father, boy friends, girl friends, teachers, etc.

The inadequacy of objective tests of isolated traits is well illustrated by a simple experiment comparing profiles made from many test scores with free case descriptions. The latter are very much easier for the individual or his friends to identify. The similarity in structure imposed by the profile actually conceals what needs to be revealed, the unique structure of pattern in each character. Such a study has been made by von Bracken (28, 1926) with much this outcome. The division of science which attempts to discover and classify these patterns or structures is called *characterology*. The types proposed by the many authors listed in the appended bibliography cannot be described and compared within the limits of this review. Roback (277, 1927) and, better, Kronfeld (273, 1932) give valuable summaries, but each takes a full book. Here we must be content with pointing out that the development of better character tests demands the laying aside, temporarily, of statistical collections of behavior reports in very limited situations, and a study of the features which make a character a character in the sense in which a portrait is a portrait and not a collection of curves and angles. Once this is better understood we may find it possible to create character tests which give us more insight into the guidance of the individual and the organization of education than can at present be claimed.

## Summary

The function of a review in the growth of a culture area should be to enable further studies to proceed from that level as a base line, building on the sound foundations, eliminating the errors of the past. We have mentioned many of the technics, possibilities, and valuable findings, so far contributed. It remains to express certain hopes for the future.

1. Character testing may be improved by better characterology. We can test almost any conceivable trait, today, but are far from testing character.

2. Better characterology will give us better units for study. The inadequacy of the ethical "trait" has been demonstrated too often to need further study. The improvement of character testing will mean testing in units which are the same in their dynamics, their inner structure, in laboratory, home, school, office, or senate.

3. Better character tests will not be content with the measurement of behavior in a few situations, but will present experimental evidence that the pattern really is identical in a wide variety of social and material environments. This is validation.

4. The reliability of tests will be improved, not so much by mere increase in length, as by more accurate insight into the behavior involved, so that errors of misinterpretation and mistaken expectation of consistency are removed.

5. The improved character tests will not attend so exclusively to something supposed to be "in" the individual, but will depend upon the inclusion of a carefully analyzed environmental setting in the behavior-definition.

6. The improved character tests will recognize character more largely as a cultural entity than as a physiological pattern, and will necessarily define the civilization in which the results are obtained and also demonstrate differences in correlation with differences in social life. The values sought and the means of seeking them will be understood to vary with the culture.

7. The improved character tests will show more contact with the life and death struggle of this generation to create the economic, political, family, and other institutions which will minister to a life of wisdom, courage, serenity, friendliness, and growth.

## Bibliography on Character Tests

1. ALLPORT, GORDON W. "Mental Tests and Measurements." *American Year Book*, 1926. New York: Macmillan Co., 1927. p. 967-71.
2. ALLPORT, GORDON W. "Personality and Character." *Psychological Bulletin* 18: 441-55; September, 1921.
3. ALLPORT, GORDON W., and VERNON, PHILIP E. "The Field of Personality." *Psychological Bulletin* 27: 677-730; December, 1930.
4. AMERICAN PSYCHIATRIC ASSOCIATION, COMMITTEE ON RELATIONS WITH THE SOCIAL SCIENCES. *Proceedings, Second Colloquium on Personality Investigation*. Baltimore: Johns Hopkins University Press, 1930. 206 p.
5. FARIS, ELLSWORTH. "Topical Summaries of Current Literature: Social Psychology in America." *American Journal of Sociology* 32: 623-30; January, 1927.
6. FROEMMING, ALBERT H. "Bibliography of Character Tests and Measurements." *Journal of Educational Research* 16: 223-26; October, 1927.
7. FURFEY, PAUL H. "Tests for Personality Traits—A Review of the Literature." *Catholic Educational Review* 25: 614-21; December, 1927.
8. FURFEY, PAUL H. *Tests for the Measurement of Non-Intellectual Traits*. Catholic University of America, Educational Research Bulletin, Vol. 3, No. 8. Washington, D. C.: Catholic Education Press, 1928. 35 p.
9. HENDRICK, IVES. "The Analysis of Personality." *American Journal of Psychiatry* 8: 535-63; November, 1928.
10. LOUTITT, CHAUNCEY M., comp. *A Bibliography of Bibliographies on Psychology, 1900-1927*. Bulletin, No. 65. Washington, D. C.: National Research Council, 1928. 108 p.
11. McDONOUGH, SISTER MARY ROSA. *The Empirical Study of Character*. Catholic University of America, Studies in Psychology and Psychiatry, Vol. 2, Nos. 3-4. Baltimore: Williams and Wilkins Co., 1929. 2 vol.
12. MANSON, GRACE E., comp. *A Bibliography of the Analysis and Measurement of Human Personality up to 1926*. Reprint and Circular Series, No. 72. Washington, D. C.: National Research Council, 1926. 59 p.
13. MAY, MARK A., and HARTSHORNE, HUGH. "Objective Methods of Measuring Character." *Pedagogical Seminary and Journal of Genetic Psychology* 32: 45-67; March, 1925.
14. MAY, MARK A., and HARTSHORNE, HUGH. "Personality and Character Tests." *Psychological Bulletin* 23: 395-411; July, 1926.
15. MAY, MARK A.; HARTSHORNE, HUGH; and WELTY, RUTH E. "Personality and Character Tests." *Psychological Bulletin* 24: 418-35; July, 1927.
16. MAY, MARK A.; HARTSHORNE, HUGH; and WELTY, RUTH E. "Personality and Character Tests." *Psychological Bulletin* 25: 422-43; July, 1928.
17. MAY, MARK A.; HARTSHORNE, HUGH; and WELTY, RUTH E. "Personality and Character Tests." *Psychological Bulletin* 26: 418-44; July, 1929.
18. MAY, MARK A.; HARTSHORNE, HUGH; and WELTY, RUTH E. "Personality and Character Tests." *Psychological Bulletin* 27: 485-94; June, 1930.
19. MONROE, WALTER S., and ASHER, OLLIE. *A Bibliography of Bibliographies*. University of Illinois Bulletin, Vol. 27, No. 44, Bureau of Educational Research Bulletin, No. 36. Urbana, Ill.; the University, 1927. 60 p.
20. NATIONAL EDUCATION ASSOCIATION, COMMITTEE ON CHARACTER EDUCATION. *Character Education*. U. S. Dept. of the Interior, Bureau of Education, Bulletin, 1926, No. 7. Washington, D. C.: Government Printing Office, 1926. Chapter 5, "Character Tests and Measurements," p. 35-48.
21. NATIONAL EDUCATION ASSOCIATION, DEPARTMENT OF SUPERINTENDENCE. "Character Education." *Tenth Yearbook*. Washington, D. C.: the Association, 1932. 536 p.
22. ROBACK, A. A. *A Bibliography of Character and Personality*. Cambridge, Mass.: Sci-Art Publishers, 1927. 340 p.
23. SHUTTLEWORTH, FRANK K. "The Social Relations of Children." *Psychological Bulletin* 24: 708-16; December, 1927.
24. SYMONDS, PERCIVAL M. "The Present Status of Character Measurement." *Journal of Educational Psychology* 15: 484-98; November, 1924.



25. THOMAS, WILLIAM I., and THOMAS, DOROTHY S. *The Child in America*. New York: Alfred A. Knopf, 1928. 583 p.
26. UPDEGRAFF, C. LESLIE. "Present Status of Moral Education in the Public Schools." *Phi Delta Kappan* 9: 137-40; April, 1927.
27. VON BRACKEN, HELMUT. "Eleven Personality Conceptions on the Basis of Descriptions of Personality." *Jena Contributions to Child and Educational Psychology, No. 1*. Jena: Psychological Institute, University of Jena, 1925.
28. WATSON, GOODWIN B. "Character Tests of 1926." *Vocational Guidance Magazine* 5: 289-309; April, 1927.
29. WATSON, GOODWIN B. *Experimentation and Measurement in Religious Education*. New York: Association Press, 1927. 295 p.
30. WATSON, GOODWIN B. "Measures of Character and Personality." To be published in the *Psychological Bulletin* during 1932.
31. WATSON, GOODWIN B. "The Measurement of the Less Tangible Qualities." *Vocational Guidance Magazine* 4: 281-89; March, 1926.
32. WATSON, GOODWIN B. "A Supplementary Review of Measures of Personality Traits." *Journal of Educational Psychology* 18: 73-87; February, 1927.
33. WATSON, GOODWIN B., and BIDDLE, DELIA H. *A Year of Research, 1927*. Religious Education Association Monograph, No. 4. Chicago: Religious Education Association, 1929. 82 p.
34. WATSON, GOODWIN B., and BIDDLE, DELIA H. *Another Year of Research*. Religious Education Association Monograph, to appear during 1932.
35. YOUNG, KIMBALL. "Topical Summaries of Current Literature: Personality Studies." *American Journal of Sociology* 32: 953-71; May, 1927.
36. YOUNG MEN'S CHRISTIAN ASSOCIATION. *Research and Studies*. Vol. 1. New York: Association Press, 1929. 75 p.
37. YOUNG MEN'S CHRISTIAN ASSOCIATION. *Research and Studies*. Vol. 2. New York: Association Press, 1930. 86 p.

#### Abnormalities, Complexes, Symptoms of Maladjustment

38. FLÜGEL, J. C., and RADCLIFFE, E. J. D. "The Pressey Cross-Out Test Compared with a Questionnaire." *British Journal of Medical Psychology* 8: 112-31; Part II, 1928.
39. GARRETT, H. E., and SCHNECK, M. R. "A Study of the Discriminative Value of the Woodworth Personal Data Sheet." *Journal of General Psychology* 1: 459-71; July-October, 1928.
40. JUNG, C. G. "Diagnostische Assoziationstudien." *Journal für Psychologie und Neurologie* 6: 1-36; 7: 1-24; 1905-06.
41. KENT, GRACE H., and ROSANOFF, A. J. *A Study of Association in Insanity*. Baltimore: Lord Baltimore Press, 1910. 142 p.
42. LAIRD, DONALD A. "A Mental Hygiene and Vocational Test." *Journal of Educational Psychology* 16: 419-22; September, 1925.
43. MATHEWS, ELLEN. "A Study of Emotional Stability in Children." *Journal of Delinquency* 8: 1-40; January, 1923.
44. PRESSEY, S. L. "A Group Scale for Investigating the Emotions." *Journal of Abnormal Psychology and Social Psychology* 16: 55-64; April, 1921.
45. ROGERS, CARL R. *Measuring Personality Adjustment in Children Nine to Thirteen Years of Age*. Contributions to Education, No. 458. New York: Teachers College, Columbia University, 1931. 107 p.
46. RORSCHACH, HERMANN. *Psychodiagnostik: Methodik und Ergebnisse eines wahrnehmungs-diagnostischen Experiments*. Berne: E. Birchner, 1921. 174 p.
47. SWEET, LENNIC. *The Measurement of Personal Attitudes in Younger Boys*. New York: Association Press, 1929. 58 p.
48. THURSTONE, L. L., and THURSTONE, THELMA G. "A Neurotic Inventory." *Journal of Social Psychology* 1: 1-29; February, 1930.
49. TRAVIS, ROLAND C. "The Diagnosis of Character Types by Visual and Auditory Thresholds." *University of Iowa Studies in Psychology, No. 10*. Psychological Monographs, Vol. 36, No. 2. Whole Number 168. Princeton, N. J.: Psychological Review Co., 1926. p. 18-37.

50. TYLER, HENRY T. *A Study of the Bearing of Certain Personality Factors Other Than Intelligence Upon Academic Success at Teachers College*. Contributions to Education, No. 468. New York: Teachers College, Columbia University, 1930. 89 p.
51. WOODROW, HERBERT H., and LOWELL, FRANCES. *Children's Association Frequency Tables*. Psychological Monographs, Vol. 22, No. 5. Whole No. 97. Princeton, N. J.: Psychological Review Co., 1916. 110 p.
52. WOODWORTH, ROBERT S., and WELLS, FREDERIC L. *Association Tests*. Psychological Monographs, Vol. 13, No. 5. Whole No. 57. Princeton, N. J.: Psychological Review Co., 1911. 85 p.

#### Accuracy

53. HARTMANN, GEORGE W. *Precision and Accuracy*. Archives of Psychology, No. 100. New York: Columbia University Press, 1928. 42 p.
54. KLINEBERG, OTTO. "Racial Differences in Speed and Accuracy." *Journal of Abnormal and Social Psychology* 22: 273-77; October-December, 1927.

#### Activity

55. BARKER, MARGARET. *A Technique for Studying the Social-Material Activities of Young Children*. Child Development Monographs, No. 3. New York: Teachers College, Columbia University, 1930. 74 p.
56. BÜHLER, C.; HETZER H.; and TUDOR-HART, B. *Soziologische und Psychologische Studien über das erste Lebensjahr*. Jena: Fischer, 1927. 250 p.
57. LOOMIS, ALICE MARIE. *A Technique for Observing the Social Behavior of Nursery School Children*. Child Development Monographs, No. 5. New York: Teachers College, Columbia University, 1931. 100 p.
58. OLSON, WILLARD C. *The Measurement of Nervous Habits in Normal Children*. Institute of Child Welfare, Monograph Series No. 3. Minneapolis: University of Minnesota Press, 1929. 97 p.
59. PRATT, KARL C.; NELSON, AMALIE K.; and SUN, KUO-HUA. *The Behavior of the Newborn Infant*. Contributions in Psychology, No. 10. Columbus: Ohio State University Press, 1930. 237 p.
60. THOMAS, DOROTHY S., and others. *Some New Techniques for Studying Social Behavior*. Child Development Monographs, No. 1. New York: Teachers College, Columbia University, 1929. 203 p.

#### Aesthetic Response

61. KARWOSKI, THEODORE F., and CHRISTENSEN, ERWIN O. "A Test for Art Appreciation." *Journal of Educational Psychology* 17: 187-94; March, 1926.
62. THORNDIKE, EDWARD L. "Tests of Esthetic Appreciation." *Journal of Educational Psychology* 7: 509-22; November, 1916.

#### Aggressiveness, Dominance, Ascendancy, Submission

63. ALLPORT, GORDON W. "A Test for Ascendancy-Submission." *Journal of Abnormal and Social Psychology* 23: 118-36; July, 1928.
64. MOORE, H. T., and GILLILAND, A. R. "The Measurement of Aggressiveness." *Journal of Applied Psychology* 5: 97-118; June, 1921.

#### Appearance as an Index of Character

65. CLEETON, GLEN U., and KNIGHT, F. B. "Validity of Character Judgments Based on External Criteria." *Journal of Applied Psychology* 8: 215-31; June, 1924.
66. DUNLAP, KNIGHT. "Reading of Character from External Signs." *Scientific Monthly* 15: 153-65; August, 1922.
67. HOLLINGWORTH, HARRY L. *Judging Human Character*. New York: D. Appleton and Co., 1922. 268 p.

#### Confidence, Inferiority

68. HEIDBREDER, EDNA F. "The Normal Inferiority Complex." *Journal of Abnormal and Social Psychology* 22: 243-58; October-December, 1927.

69. LUND, FREDERICK H. "The Criteria of Confidence." *American Journal of Psychology* 37: 372-81; July, 1926.
70. TROW, W. C. *The Psychology of Confidence*. Archives of Psychology, No. 67. New York: Columbia University Press, 1923. 47 p.

#### Cooperation, Service, Negativism

71. HARTSHORNE, HUGH; MAY, MARK A.; and MALLER, JULIUS B. *Studies in Service and Self-Control*. New York: Macmillan Co., 1929. 559 p.
72. MALLER, JULIUS B. *Cooperation and Competition*. Contributions to Education, No. 384. New York: Teachers College, Columbia University, 1929. 176 p.
73. NELSON, JANET FOWLER. *Personality and Intelligence*. Child Development Monographs, No. 4. New York: Teachers College, Columbia University, 1931. 62 p.
74. REYNOLDS, MARTHA M. *Negativism of Pre-School Children*. Contributions to Education, No. 288. New York: Teachers College, Columbia University, 1928. 126 p.
75. RUST, METTA MAUND. *The Effect of Resistance on Intelligence Test Scores of Young Children*. Child Development Monographs, No. 6. New York: Teachers College, Columbia University, 1931. 80 p.
76. ZALUZHNY, A. S. *The Character of the Interaction between Children of Pre-School Age*. Ukarinski Vestnik Eksperimentalnoe Pegagogiki i Refleksologii, No. 2, 1927.

#### Delinquent Trends

77. BURT, CYRIL L. *The Young Delinquent*. New York: D. Appleton and Co., 1925. 619 p.
78. CADY, VERNON M. *The Estimation of Juvenile Incurability*. Journal of Delinquency Monograph No. 2. Whittier, Calif.: California Bureau of Juvenile Research, Whittier State School, 1923. 140 p.
79. CUSHING, HAZEL M., and RUCH, G. M. "An Investigation of Character Traits in Delinquent Girls." *Journal of Applied Psychology* 11: 1-7; February, 1927.
80. GLUECK, SHELDON, and GLUECK, ELEANOR T. *500 Criminal Careers*. New York: Alfred A. Knopf, 1930. 365 p.
81. HEALY, WILLIAM, and BRONNER, AUGUSTA F. *Delinquents and Criminals, Their Making and Unmaking*. New York: Macmillan Co., 1926. 317 p.
82. LENTZ, THEODORE F. *An Experimental Method for the Discovery and Development of Tests of Character*. Contributions to Education, No. 180. New York: Teachers College, Columbia University, 1925. 47 p.

#### Emotions

83. GAUGER, MARGUERITE. *Modifiability of Response to Taste Stimuli in Pre-School Children*. Contributions to Education, No. 348. New York: Teachers College, Columbia University, 1929. 53 p.
84. HERRING, JOHN P. "The Measurement of Liking and Disliking." *Journal of Educational Psychology* 21: 159-96; March, 1930.
85. JONES, H. E., and JONES, M. C. "Genetic Studies of Emotions." *Psychological Bulletin* 27: 40-64; January, 1930.
86. LANDIS, CARNEY; GULLETTE, RUTH; and JACOBSON, CARLYLE. "Criteria of Emotionality." *Pedagogical Seminary and Journal of Genetic Psychology* 32: 209-34; June, 1925.
87. SMITH, W. WHATELY. *The Measurement of Emotion*. New York: Harcourt, Brace and Co., 1922. 183 p.
88. STRATTON, GEORGE M. "Emotion and the Incidence of Disease." *Journal of Abnormal and Social Psychology* 21: 19-23; April, 1926.
89. WASHBURN, MARGARET FLOY. "Feeling and Emotion." *Psychological Bulletin* 24: 573-95; October, 1927.

#### Excitability

90. CASON, HULSEY. *Common Annoyances*. Psychological Monographs, Vol. 40, No. 2. Whole No. 182. Princeton, N. J.: Psychological Review Co., 1930. 218 p.
91. RICH, GILBERT J. "Body Acidity as Related to Emotional Excitability." *Archives of Neurology and Psychiatry* 20: 589-94; September, 1923.

### Expression, Handwriting, Will-Temperament

92. BINET, ALFRED. *Les Revelations de l'Ecriture d'après un Controle Scientifique*. Paris: F. Alcan, 1906. 260 p.
93. DOWNEY, JUNE E. *The Will-Temperament and Its Testing*. Yonkers-on-Hudson, N. Y.: World Book Co., 1923. 339 p.
94. HULL, CLARK L., and MONTGOMERY, ROBERT B. "An Experimental Investigation of Certain Alleged Relations between Character and Handwriting." *Psychological Review* 26: 63-74; January, 1919.
95. KLAGES, LUDWIG. *Handschrift und Charakter*. Leipzig: J. A. Barth, 1920. 254 p.
96. MAY, MARK A. "The Present Status of the Will-Temperament Tests." *Journal of Applied Psychology* 9: 29-52; March, 1925.
97. UHRBROCK, RICHARD S. *An Analysis of the Downey Will-Temperament Tests*. Contributions to Education, No. 296. New York: Teachers College, Columbia University, 1928. 78 p.
98. WAGONER, LOVISA C., and DOWNEY, JUNE E. "Speech and Will-Temperament." *Journal of Applied Psychology* 6: 291-97; September, 1922.
99. WOLFF, WERNER. "Ueber Faktoren Charakterologischen Urteilsbildung." *Zeitschrift für Angewandte Psychologie* 35: 385-446; 1930.

### Happiness, Cheerfulness

100. ENDERS, ABBIE CRANDALL. "A Study of the Laughter of the Pre-School Child in the Merrill-Palmer Nursery School." *Papers of the Michigan Academy of Science, Art, and Letters*. Vol. 8. New York: Macmillan Co., 1928. p. 341-56.
101. HAMILTON, GILBERT VAN T. *A Research in Marriage*. New York: Albert and Charles Boni, 1929. 570 p.
102. MORGAN, ELEANOR; MULL, HELEN K.; and WASHBURN, M. F. "An Attempt to Test Moods or Temperaments of Cheerfulness and Depression by Directed Recall of Emotionally Toned Experiences." *American Journal of Psychology* 30: 302-4; June, 1919.
103. SAILER, RANDOLPH C. *Happiness Self-Estimates of Young Men*. Contributions to Education, No. 467. New York: Teachers College, Columbia University, 1931. 116 p.
104. WATSON, GOODWIN B. "Happiness Among Adult Students of Education." *Journal of Educational Psychology* 21: 79-109; February, 1930.

### Home Background

105. CHAPMAN, J. CROSBY, and SIMS, V. M. "The Quantitative Measurement of Certain Aspects of Socio-Economic Status." *Journal of Educational Psychology* 16: 380-90; September, 1925.

### Honesty, Deception

106. ADLER, HERMAN, and LARSON, JOHN A. "Deception and Self-Deception." *Journal of Abnormal and Social Psychology* 22: 364-71; January-March, 1928.
107. CROSLAND, HAROLD R. *The Psychological Methods of Word-Association and Reaction-Time as Tests of Deception*. University of Oregon Publication, Psychology Series, Vol. 1. No. 1. Eugene, Ore.: the University, 1929. 104 p.
108. HARTSHORNE, HUGH, and MAY, MARK A. *Studies in Deceit*. New York: Macmillan Co., 1928. Book I, 414 p.; Book II, 306 p.
109. LANDIS, CARNEY. "Detecting the Deceiver." *Industrial Psychology* 2: 244-49; May, 1927.
110. VOELKER, PAUL F. *The Function of Ideals and Attitudes in Social Education*. Contributions to Education, No. 112. New York: Teachers College, Columbia University, 1921. 126 p.
111. WATSON, GOODWIN B. *Some Accomplishments in Summer Camps*. New York: Association Press, 1928. 28 p.
112. WOODROW, HERBERT, and BEMMELS, VIOLA. "Overstatement as a Test of General Character in Pre-School Children." *Journal of Educational Psychology* 18: 239-46; April, 1927.

### Humor

- 113. BIRD, GRACE E. "An Objective Humor Test for Children." *Psychological Bulletin* 22: 137-38; February, 1925.
- 114. DISERENS, CHARLES M., and BONIFIELD, MABEL. "Humor and the Ludicrous." *Psychological Bulletin* 27: 108-18; February, 1930.
- 115. KAMBOUROPOULOU, POLYXENIE. "Individual Differences in the Sense of Humor." *American Journal of Psychology* 37: 268-78; April, 1926.

### Inhibition, Caution, Self-Control

- 116. CRANE, ALBERT L. *Race Differences in Inhibition*. Archives of Psychology, No. 63. New York: Columbia University Press, 1923. 86 p.
- 117. SNOW, A. J. "Tests for Transportation Pilots." *Journal of Applied Psychology* 10: 37-51; March, 1926.
- 118. WASHBURN, JOHN N. "An Experiment in Character Measurement." *Journal of Juvenile Research* 13: 1-18; January, 1929.

### Interests

- 119. BRIDGES, K. M. BANHAM. "Occupational Interests and Attention of Four-Year-Old Children." *Pedagogical Seminary and Journal of Genetic Psychology* 36: 551-69; December, 1929.
- 120. CARNEGIE INSTITUTE OF TECHNOLOGY, BUREAU OF PERSONNEL RESEARCH. *Range of Interest Test V, Form 803*. Pittsburgh: Carnegie Institute of Technology, 1917.
- 121. JORDAN, ARTHUR M. *Children's Interests in Reading*; rev. ed. Chapel Hill, N. C.: University of North Carolina Press, 1926. 103 p.
- 122. LANDIS, CARNEY. "National Differences in Conversation." *Journal of Abnormal and Social Psychology* 21: 354-57; January, 1927.
- 123. LEHMAN, HARVEY C., and WITTY, PAUL A. *The Psychology of Play Activities*. New York: A. S. Barnes and Co., 1927. 242 p.
- 124. McHALE, KATHRYN. "An Experimental Study of the Vocational Interests of a Liberal Arts College Group." *Journal of Applied Psychology* 8: 245-55; June, 1924.
- 125. MINER, J. B. "A Method for Evaluating a Psychograph for Vocational Guidance." *Journal of Educational Psychology* 17: 331-40; May, 1926.
- 126. STRONG, EDWARD K. "A Vocational Interest Test." *Educational Record* 8: 107-21; April, 1927.
- 127. Terman, Lewis M., and Lima, Margaret. *Children's Reading*. New York: D. Appleton and Co., 1926. 363 p. (Second edition published in 1931; 422 p.)
- 128. UHRBROCK, R. S. "Interest as an Indication of Ability." *Journal of Applied Psychology* 10: 487-501; December, 1926.
- 129. WYMAN, JENNIE B. "The Measurement of Interest." *Vocational Guidance Magazine* 8: 54-60; November, 1929.

### Introversion, Extroversion

- 130. FREYD, MAX. "Introverts and Extroverts." *Psychological Review* 31: 74-87; January, 1924.
- 131. GUILFORD, J. P., and BRALY, KENNETH W. "Extroversion and Introversion." *Psychological Bulletin* 27: 96-107; February, 1930.
- 132. GUTHRIE, EDWIN R. "Measuring Introversion and Extroversion." *Journal of Abnormal and Social Psychology* 22: 82-88; April, 1927.
- 133. MARSTON, L. R. *The Emotions of Young Children*. University of Iowa Studies in Child Welfare, Vol. 3, No. 3. Iowa City; the University, 1925. 99 p.
- 134. NEWCOMB, THEODORE M. *The Consistency of Certain Extrovert-Introvert Behavior Patterns in Fifty-One Problem Boys*. Contributions to Education, No. 382. New York: Teachers College, Columbia University, 1929. 123 p.

### Leadership

- 135. BOWDEN, ABERDEEN O. "A Study of the Personality of Student Leaders in Colleges in the United States." *Journal of Abnormal and Social Psychology* 21: 149-60; July, 1926.



136. GOODENOUGH, FLORENCE L. "Measuring Behavior Traits by Means of Repeated Short Samples." *Journal of Juvenile Research* 12: 230-35; September-December, 1928.
137. MOXCEY, MARY E. *Some Qualities Associated with Success in the Christian Ministry*. Contributions to Education, No. 122. New York: Teachers College, Columbia University, 1922. 101 p.
138. Terman, Lewis M., and others. *Genetic Studies of Genius*. Vols. 1-3. Stanford University, Calif.: Stanford University Press, 1925-30.

#### Maturity: Social and Emotional

139. FURFEY, PAUL H. "Developmental Age." *American Journal of Psychiatry* 8: 149-57; July, 1928.
140. GATES, ARTHUR I. "The Nature and Educational Significance of Physical Status and of Mental, Physiological, Social, and Emotional Maturity." *Journal of Educational Psychology* 15: 329-58; September, 1924.
141. GATES, GEORGINA S. "An Experimental Study of the Growth of Social Perception." *Journal of Educational Psychology* 14: 449-61; November, 1923.

#### Moral Knowledge, Ethical Judgment

142. ATHEARN, W. S., and others. *Measurements and Standards in Religious Education*. Indiana Survey of Religious Education, Vol. 2. New York: George H. Doran Co. (now Doubleday, Doran and Co.), 1924. 532 p.
143. BROTEMARKLE, R. A. "A Comparison Test for Investigating the Ideational Content of the Moral Concepts." *Journal of Applied Psychology* 6: 235-42; September, 1922.
144. CHASSELL, CLARA F.; CHASSELL, ELLA B.; and CHASSELL, LAURA M. "A Test of Ability to Weigh Foreseen Consequences." *Teachers College Record* 25: 39-50; January, 1924.
145. FERNALD, GUY G. "The Defective Delinquent Class Differentiating Tests." *American Journal of Insanity* 68: 523-94; April, 1912.
146. HARTSHORNE, HUGH; MAY, MARK A.; and SHUTTLEWORTH, FRANK K. *Studies in the Organization of Character*. New York: Macmillan Co., 1930. 503 p.
147. KOHS, S. C. "Ethical Discrimination Test." *Journal of Delinquency* 7: 1-15; January, 1922.
148. LINCOLN, EDWARD A., and SHIELDS, FRED J. "An Age Scale for the Measurement of Moral Judgment." *Journal of Educational Research* 23: 193-97; March, 1931.
149. LOWE, GLADYS M., and SHIMBERG, MYRA E. "A Critique of the Fables as a Moral Judgment Test." *Journal of Applied Psychology* 9: 53-59; March, 1925.
150. MACAULAY, E., and WATKINS, STANLEY H. "An Investigation into the Development of the Moral Conceptions of Children." *Forum of Education* 4: 13-33, 92-108; February and June, 1926.
151. MAY, MARK A. "The Psychological Examination of Conscientious Objectors." *American Journal of Psychology* 31: 152-65; April, 1920.
152. SCHWESINGER, GLADYS C. *The Social-Ethical Significance of Vocabulary*. Contributions to Education, No. 211. New York: Teachers College, Columbia University, 1926. 73 p.
153. SHARP, FRANK CHAPMAN. "An Objective Study of Some Moral Judgments." *American Journal of Psychology* 9: 198-234; January, 1898.
154. WATSON, GOODWIN B. *Some Experiments with Religious Education Tests*. Program Paper No. 5. New York: Association Press, 1926. 50 p.
155. WOODROW, HERBERT H. "A Picture-Preference Character Test." *Journal of Educational Psychology* 17: 519-31; November, 1926.

#### Morphology, Constitutional Type, Physical Build

156. KRASUSKY, W. S. "Kretschmer's Konstitutionelle Typen unter den Kindern im Schulalter." *Archiv für Kinderheilkunde* 82: 22-32; 1927.
157. KRETSCHEMER, ERNST. *Körperbau und Charakter*. Berlin: J. Springer, 1922. 195 p. (English translation published in 1925 by Harcourt, Brace and Co., New York.)

158. NACCARATI, SANTE, and GARRETT, HENRY E. "The Influence of Constitutional Factors in Behavior." *Journal of Experimental Psychology* 6: 455-65; December, 1923.
159. NACCARATI, SANTE, and GARRETT, HENRY E. "The Relation of Morphology to Temperament." *Journal of Abnormal and Social Psychology* 19: 254-63; October, 1924.
160. PATERSON, DONALD. *Physique and Intellect*. New York: Century Co., 1930. 304 p.
161. POLEN, L. "Körperbau und Charakter." *Archiv für die gesamte Psychologie* 66: 1-116; 1928.
162. SHELDON, WILLIAM H. "Social Traits and Morphologic Types." *Personnel Journal* 6: 47-55; June, 1927.

#### Opinions, Attitudes, Prejudices

163. ALLPORT, FLOYD H., and HARTMAN, D. A. "Measurement and Motivation of Atypical Opinion in a Certain Group." *American Political Science Review* 19: 735-60; November, 1925.
164. BAIN, READ. "Theory and Measurement of Attitudes and Opinions." *Psychological Bulletin* 27: 357-79; May, 1930.
165. BOCARDUS, EMORY S. "Measuring Social Distances." *Journal of Applied Sociology* 9: 299-308; March, 1925.
166. DAVIS, KATHARINE BEMENT. *Factors in the Sex Life of Twenty-Two Hundred Women*. New York: Harper and Brothers, 1929. 430 p.
167. ELLIOTT, HARRISON SACKETT. *The Process of Group Thinking*. New York: Association Press, 1928. 229 p.
168. HARPER, HEBER REECE. *What European and American Students Think on International Problems*. New York: Teachers College, Columbia University, 1931. 255 p.
169. HARPER, MANLEY H. *Social Beliefs and Attitudes of American Educators*. Contributions to Education, No. 294. New York: Teachers College, Columbia University, 1927. 91 p.
170. LEONARD, EUGENIE ANDREWS. *Concerning Our Girls and What They Tell Us*. Contributions to Education, No. 430. New York: Teachers College, Columbia University, 1930. 192 p.
171. MACLEAN, ANGUS HECTOR. *The Idea of God in Protestant Religious Education*. Contributions to Education, No. 410. New York: Teachers College, Columbia University, 1930. 150 p.
172. MOORE, HENRY T. "Innate Factors in Radicalism and Conservatism." *Journal of Abnormal and Social Psychology* 20: 234-44; October, 1925.
173. RICE, S. A. *Quantitative Methods in Politics*. New York: Alfred A. Knopf, 1928. 331 p.
174. THURSTONE, L. L., and CHAVE, E. J. *The Measurement of Attitude*. Chicago: University of Chicago Press, 1929. 96 p.
175. VETTER, GEORGE B. "The Study of Social and Political Opinions." *Journal of Abnormal and Social Psychology* 25: 26-39; April, 1930.
176. WATSON, GOODWIN B. *Orient and Occident; an Opinion Study*. New York: Institute of Pacific Relations, 1927. 120 p.
177. WATSON, GOODWIN B. *The Measurement of Fair-mindedness*. Contributions to Education, No. 176. New York: Teachers College, Columbia University, 1925. 97 p.

#### Originality, Imagination, Resourcefulness

178. CHASSELL, LAURA M. "Tests for Originality." *Journal of Educational Psychology* 7: 317-28; June, 1916.
179. McCLATCHY, VIVIENNE R. "A Theoretical and Statistical Study of the Personality Trait Originality as Herein Defined." *Journal of Abnormal and Social Psychology* 23: 379-82; October, 1928.
180. MCGEOCH, J. A. "The Relationships Between Three Tests of Imagination and Their Correlation with Intelligence." *Journal of Applied Psychology* 8: 439-43; December, 1924.

### Perseveration

181. CUSHING, HAZEL M. *A Perseverative Tendency in Pre-School Children*. Archives of Psychology, No. 108. New York: Columbia University Press, 1929. 55 p.
182. LANKES, W. "Perseveration." *British Journal of Psychology* 7: 387-419; March, 1915.
183. SPEARMAN, CHARLES. *The Abilities of Man*. New York: Macmillan Co., 1927. 415 p.
184. WEBB, EDWARD. *Character and Intelligence*. British Journal of Psychology, Monograph Supplements, Vol. I, Part 3. Cambridge: Cambridge University Press, 1915. 99 p.

### Persistence, Perseverance, Effort

185. CHAPMAN, J. CROSBY. "Persistence, Success, and Speed in a Mental Task." *Pedagogical Seminary* 31: 276-84; September, 1924.
186. FERNALD, GUY G. "An Achievement Capacity Test." *Journal of Educational Psychology* 3: 331-36; June, 1912.
187. MORGAN, JOHN J. B., and HULL, HAZEL L. "The Measurement of Persistence." *Journal of Applied Psychology* 10: 180-87; June, 1926.

### Physiological Indices of Character

188. PROESCHER, F., and ARKUSH, A. S. "Blood Groups in Mental Disease." *Journal of Nervous and Mental Disease* 65: 569-84; June, 1927.
189. STRATTON, GEORGE M. "Emotion and the Incidence of Disease." *Journal of Abnormal and Social Psychology* 21: 19-23; April-June, 1926.

### Psychogalvanic Responses

190. BAYLEY, NANCY. "A Study of Fear by Means of the Psychogalvanic Technique." *University of Iowa Studies in Psychology, No. 11*. Psychological Monographs, Vol. 38, No. 4. Whole No. 176. Princeton, N. J.: Psychological Review Co., 1928. p. 1-38.
191. PRIDEAUX, E. "The Psychogalvanic Reflex; a Review." *Brain* 43: 50-73; May, 1920.
192. RICHTER, CURT P. "A Study of the Electrical Skin Resistance and the Psychogalvanic Reflex in a Case of Unilateral Sweating." *Brain* 50: 216-35; June, 1927.
193. VERACUTH, OTTO. *Das Psycho-galvanischen Reflex Phänomen*. Berlin: Krager, 1908.
194. WOODROW, HERBERT. "The Psychogalvanic Reflex." *Psychological Bulletin* 23: 237-46; May, 1926.

### Ratings, Reputation Measures

195. BERNE, ESTER VAN C. *An Experimental Investigation of Social Behavior Patterns in Young Children*. University of Iowa Studies in Child Welfare, Vol. 4, No. 3. Iowa City: the University, 1930. 93 p.
196. DIMOCK, HEDLEY S., and HENDRY, CHARLES E. *Camping and Character*. New York: Association Press, 1929. 364 p.
197. HUGHES, W. HARDIN. "General Principles and Results of Rating Trait Characteristics." *Journal of Educational Method* 4: 421-31; June, 1925.
198. KINGSBURY, FORREST A. "Making Rating-Scales Work." *Journal of Personnel Research* 4: 1-6; May, 1925.
199. KNIGHT, FREDERIC V., and FRANZEN, RAYMOND H. "Pitfalls in Rating Schemes." *Journal of Educational Psychology* 13: 204-13; April, 1922.
200. KORNHAUSER, ARTHUR W. "A Comparison of Raters." *Journal of Personnel Research* 5: 338-44; January, 1927.
201. KORNHAUSER, ARTHUR W. "A Comparison of Ratings on Different Traits." *Journal of Personnel Research* 5: 440-46; March, 1927.
202. NORSWORTHY, NAOMI. "The Validity of Judgments of Character." *Essays Philosophical and Psychological in Honor of William James*. New York: Longmans, Green and Co., 1908. p. 553-67.
203. PATERSON, DONALD. "Methods of Rating Human Qualities." *Annals* 110: 81-93; November, 1923. Philadelphia: American Academy of Political and Social Science.

204. ROBERTSON, D. A., chairman. "Report of Sub-Committee on Personality Measurement." *Educational Record*, Supplement No. 8, July, 1928. Washington, D. C.: American Council on Education. p. 53-68.
205. RUGG, HAROLD O. "Is the Rating of Human Character Practicable?" *Journal of Educational Psychology* 12: 425-38, 485-501, November-December, 1921; 13: 30-42, 81-93, January-February, 1922.
206. SHEN, EUGENE. "The Reliability Coefficient of Personal Ratings." *Journal of Educational Psychology* 16: 232-36; April, 1925.
207. STALNAKER, JOHN M., and REMMERS, H. H. "Can Students Discriminate Traits Associated with Success in Teaching?" *Journal of Applied Psychology* 12: 602-10; December, 1928.
208. THORNDIKE, EDWARD L. "Fundamental Theorems in Judging Men." *Journal of Applied Psychology* 2: 67-76; March, 1918.
209. WATSON, GOODWIN B. *Rating Scales*, 1928. New York: Association Press, 1928.
210. WICKMAN, E. KOSTER. *Children's Behavior and Teachers' Attitudes*. New York: Commonwealth Fund, 1928. 247 p.

#### School Success and Failure: Character Factors

211. CHAUNCEY, MARLIN R. "Relation of the Home Factor to Achievement and Intelligence Test Scores." *Journal of Educational Research* 20: 88-90; September, 1929.
212. HERRIOTT, M. E. *Attitudes as Factors in Scholastic Success*. University of Illinois Bulletin, Vol. 27, No. 2, Bureau of Educational Research Bulletin, No. 47. Urbana, Ill.: the University, 1929. 72 p.
213. HURLOCK, ELIZABETH B. *The Value of Praise and Reproof as Incentives for Children*. Archives of Psychology, No. 71. New York: Columbia University Press, 1924. 78 p.
214. MAY, MARK A. "Predicting Academic Success." *Journal of Educational Psychology* 14: 429-40; October, 1923.
215. REMMERS, HERMAN H. *A Diagnostic and Remedial Study of Potentially and Actually Failing Students at Purdue University*. Purdue University Bulletin, Vol. 28, No. 12. Lafayette, Ind.: the University, 1928. 164 p.
216. SHUTTLEWORTH, FRANK K. *The Measurement of Character and Environmental Factors Involved in Scholastic Success*. University of Iowa Studies in Character, Vol. I, No. 2. Iowa City: the University, 1927. 80 p.
217. SYMONDS, PERCIVAL M. "A Studiousness Questionnaire." *Journal of Educational Psychology* 19: 152-67; March, 1928.
218. TOOPS, HERBERT A. "The Prediction of Scholastic Success in College." *School and Society* 25: 265-68; February 26, 1927.
219. WATSON, GLADYS HIPPLE. *Success and Failure in the Teaching Profession*. New York: Teachers College, Columbia University, 1932. In press.

#### Self-Appraisal

220. FRANZEN, RAYMOND. "The Measurement of Attitude by Means of Self-Estimation." *Psychological Bulletin* 21: 195; April, 1924.
221. HOFFMAN, GEORGE J. "An Experiment in Self-Estimation." *Journal of Abnormal and Social Psychology* 18: 43-49; April-June, 1923.
222. SCHUTTE, T. H. "Students' Estimates of Their Ability and Achievement." *Pedagogical Seminary and Journal of Genetic Psychology* 35: 627-29; December, 1928.
223. SHAW, ROBERT W. *Some Aspects of Self-Insight as Found in Students of a Two-Year Normal School*. Contributions to Education, No. 448. New York: Teachers College, Columbia University, 1931. 78 p.
224. SHEN, EUGENE. "The Validity of Self-Estimate." *Journal of Educational Psychology* 16: 104-7; February, 1925.
225. UHRBROCK, RICHARD S. "Personal Estimates of Character Traits." *Pedagogical Seminary and Journal of Genetic Psychology* 33: 491-96; September, 1926.

### Sex Differences

226. ALLEN, C. N. "Studies in Sex Differences." *Psychological Bulletin* 24: 294-304; May, 1927.
227. ALLEN, C. N. "Recent Studies in Sex Differences." *Psychological Bulletin* 27: 394-407; May, 1930.
228. HEIDBREDER, EDNA F. "Introversion and Extroversion in Men and Women." *Journal of Abnormal and Social Psychology* 22: 52-61; April-June, 1927.
229. LIPMANN, OTTO. *Psychische Geschlechtsunterscheide. Ergebnisse die differentielle Psychologie*, Vol. 2. Leipzig: 1924.
230. MILES, CATHARINE C., and TERMAN, L. M. "Sex Difference in the Association of Ideas." *American Journal of Psychology* 41: 165-206; April, 1929.
231. VALENTINE, C. W. "The Relative Reliability of Men and Women in Intuitive Judgments of Character." *British Journal of Psychology* (General Section) 19: 213-38; January, 1929.

### Sociability, Social Acceptability

232. CHAPIN, F. STUART. "Measuring the Volume of Social Stimuli: A Study in Social Psychology." *Social Forces* 4: 479-95; March, 1926.
233. FURFEY, PAUL H. "Some Factors Influencing the Selection of Boys' Chums." *Journal of Applied Psychology* 11: 47-51; February, 1927.
234. GILLILAND, A. R., and BURKE, RUTH S. "A Measurement of Sociability." *Journal of Applied Psychology* 10: 315-26; September, 1926.
235. MOSS, F. A., and others. *Social Intelligence Test*. Washington, D. C.: Center for Psychological Service, George Washington University, 1926.
236. RASEY, MARIE, and others. *How Children Choose Friends*. Detroit, Mich.: Detroit Teachers College, 1929. 52 p.
237. SHUTTLEWORTH, FRANK K. "The Social Relations of Children." *Psychological Bulletin* 24: 708-16; December, 1927.
238. VERRY, ETHEL E. "Study of Personality in Pre-School Play Groups." *Journal of Social Forces* 3: 645-48; May, 1925.
239. ZOEFFEL, H. "Ein Versuch zur Experimentellen Feststellung der Persönlichkeit im Säuglingsalter." *Zeitschrift für Psychologie* 111: 273-306; 1929.

### Speed

240. BAXTER, MILDRED F. "An Experimental Study of the Differentiation of Temperaments on a Basis of Rate and Strength." *American Journal of Psychology* 38: 59-96; January, 1927.
241. BERNSTEIN, E. *Quickness and Intelligence; An Enquiry Concerning the Existence of a General Speed Factor*. British Journal of Psychology, Monograph Supplements, No. 7. Cambridge: Cambridge University Press, 1924. 55 p.
242. BRAUN, FELIX. "Untersuchung über das persönliche Tempo." *Archiv für die gesamte Psychologie* 60: 317-60; 1927.
243. BRIDGES, JAMES WINFRED. *An Experimental Study of Decision Types and Their Mental Correlates*. Psychological Monographs, Vol. 17, No. 1. Princeton, N. J.: Psychological Review Co., 1914. p. 51-55.
244. KENNEDY, MARGARET. "Speed as a Personality Trait." *Journal of Social Psychology* 1: 286-98; May, 1930.
245. KLINEBERG, OTTO. *Experimental Study of Speed and Other Factors in Racial Differences*. Archives of Psychology, No. 93. New York: Columbia University Press, 1928. 111 p.
246. TROW, WILLIAM C. "Trait Consistency and Speed of Decision." *School and Society* 21: 538-42; May 2, 1925.

### Suggestibility

247. BROWN, WARNER. *Individual and Sex Differences in Suggestibility*. University of California Publications in Psychology, Vol. 2, No. 6. Berkeley, Calif.: University of California Press, 1916. p. 291-430.



248. HOWELLS, THOMAS HENRY. *Comparative Study of Those Who Accept as against Those Who Reject Religious Authority*. University of Iowa Studies in Character, Vol. 2, No. 2. Iowa City: the University, 1928. 80 p.
249. MCGEOCH, JOHN A. "The Relationship between Suggestibility and Intelligence in Delinquents." *Psychological Clinic* 16: 133-34; March-April, 1925.
250. MURPHY, GARDNER, and MURPHY, L. B. *Experimental Social Psychology*. New York: Harper and Brothers, 1931. 709 p.
251. OTIS, MARGARET. *A Study of Suggestibility of Children*. Archives of Psychology, No. 70. New York: Columbia University Press, 1924. 109 p.
252. SINCLAIR, ROBERT DANIEL. *Comparative Study of Those Who Report the Experience of the Divine Presence and Those Who Do Not*. University of Iowa Studies in Character, Vol. 2, No. 3. Iowa City: the University, 1930. 63 p.
253. WHITTEMORE, IRVING C. "Competitive Consciousness." *Journal of Abnormal and Social Psychology* 20: 17-33; April, 1925.

### Types: Underlying Organization of Character

254. ACHILLE-DELMAS, F., and BOLL, N. *La Personnalité Humaine; son Analyse*. Paris: E. Flammarion, 1922. 283 p.
255. ADLER, ALFRED. *Understanding Human Nature*. New York: Greenberg, Publisher, 1927. 286 p.
256. ALEXANDER, FRANZ. *Psychoanalyse der Gesamtpersönlichkeit*. Leipzig: International Psycho-Analytische Verlag, 1927. 240 p. (English translation published in 1930 by the Nervous and Mental Disease Publishing Co., New York.)
257. BURT, CYRIL. "The Mental Differences Between Individuals." *Report*, 1923. British Association for Advancement of Science. London: the Association, 1924. p. 215-39.
258. EWALD, G. *Temperament und Charakter*. Berlin: J. Springer, 1924. 156 p.
259. FRANK, LAWRENCE K. "The Management of Tensions." *American Journal of Sociology* 33: 705-36; March, 1928.
260. FREUD, SIGMUND. *Gesammelte Schriften, Vol. 12*. Vienna: International Psycho-Analytische Verlag, through 1932.
261. HAEERLIN, PAUL. *Der Charakter*. Basel: 1925.
262. HEYMANS, G. D., and WIERSMA, E. "Beiträge zur speziellen Psychologie auf Grund einer Massenuntersuchung." *Zeitschrift für Psychologie* 42:81-127; 258-301; 1906.
263. HINKLE, BEATRICE M. "A Study of Psychological Types." *Psycho-analytic Review* 9: 107-97; April, 1922.
264. HOFFMAN, H. *Das Problem des Charakteraufbaus*. Berlin: J. Springer, 1926. 193 p.
265. JAENSCH, ERICH. *Eidetic Imagery and Typological Methods of Investigation*. London: George Routledge and Sons, Ltd., 1930. 136 p.
266. JAENSCH, E. R. *Mitarbeitern, Grundformen Menschlichen Seins*. Berlin: 1929.
267. JAENSCH, E. R. *Über Mitarbeitern Studien zur Psychologie Menschlicher Typen*. Leipzig: Barth, 1930. 532 p.
268. JONES, ERNEST. *Papers on Psycho-analysis; 3d rev. ed.* New York: William Wood and Co., 1923.
269. KELLEY, TRUMAN L. "Oddities in Mental Make-up." *School and Society* 24: 529-35; October 30, 1926.
270. KLAGES, LUDWIG. *Prinzipien der Charakterologie*. Leipzig: J. Barth, 1923. 95 p.
271. KLAGES, LUDWIG. *The Science of Character*. London: George Allen and Unwin, Ltd., 1929. 308 p.
272. KLÜVER, HEINRICH. "An Analysis of Recent Work on the Problem of Psychological Types." *Journal of Nervous and Mental Disease* 62: 561-96; December, 1925.
273. KRONFELD, ARTHUR. *Lehrbuch der Charakterkunde*. Berlin: J. Springer, 1932. 451 p.
274. KÜNKEL, F. *Einführung in die Charakterkunde auf individual-psychologischer Grundlage*. Leipzig: Hirzel, 1928. 185 p.
275. LEWIN, K., and others. "Untersuchungen zur Handlungs-und Affekt-Psychologie." *Psychologische Forschung*. Vol. 7 through 13; 1926-32. A series of studies.

276. PAVLOV, IVAN P. *Lectures on Conditioned Reflexes*. New York: International Publishers Co., 1928. 414 p.
277. ROBACK, A. A. *The Psychology of Character*. New York: Harcourt, Brace and Co., 1927. 595 p.
278. SELZ, OTTO. *Über die Persönlichkeitstypen und die Methoden ihrer Bestimmung*. Jena: Fischer, 1924. 44 p.
279. SHAW, CLIFFORD R. "Case Study Method." *Proceedings*, 1926. Vol. 21. American Sociological Society. Chicago: University of Chicago Press, 1927. p. 149-57.
280. SHAND, ALEXANDER F. *The Foundations of Character*; 2d ed. New York: Macmillan Co., 1920. 578 p.
281. SPRANGER, EDUARD. *Types of Men*. Halle (Salle): M. Niemeyer, 1928. 402 p.
282. THOMAS, W. I. "The Configurations of Personality." *The Unconscious; a Symposium*. New York: Alfred A. Knopf, 1928. p. 143-77.